

Universidade do Minho
Escola de Engenharia
Departamento de Informática

Projecto e Implementação de Sistemas de *Data Warehousing*

Vasco Nuno Caio dos Santos

Dissertação de Mestrado

2004

Projecto e Implementação de Sistemas de *Data Warehousing*

Vasco Nuno Caio dos Santos

Dissertação apresentada à Universidade do Minho para obtenção do grau de Mestre em Informática,
na especialidade de Sistemas Distribuídos, Comunicações por Computador e Arquitectura de
Computadores, elaborada sob orientação do Professor Doutor Orlando Manuel de Oliveira Belo.

2004

Aos que me querem bem.

Agradecimentos

Gostaria de agradecer a todos aqueles que me apoiaram e contribuíram para a conclusão deste Mestrado, e em particular :

- Ao Professor Doutor Orlando Belo pela orientação, acompanhamento, incentivo e permanente disponibilidade demonstrada durante este projecto.
- Ao Professor Doutor Alexandre Rodrigues pelo apoio demonstrado e pelas suas críticas e sugestões no estudo da área de Gestão de Projectos.
- Aos meus colegas de Mestrado que através do companheirismo e entreaajuda tornaram mais fácil a conclusão do ano curricular.
- À Direcção da Escola Superior de Tecnologia e Gestão de Felgueiras - ESTGF pelas condições oferecidas e pelo apoio concedido necessários à conclusão deste projecto.
- Aos colegas de trabalho da ESTGF que me apoiaram e incentivaram durante este longo caminho.
- À minha família, e em particular aos meus pais, pelo permanente estímulo e natural apoio demonstrado ao longo da realização desta dissertação e trabalhos relacionados.

Resumo

Nos últimos anos é comum ouvir-se dizer que se está perante a *Era da Informação*. A massificação das Tecnologias da Informação contribuiu de forma muito significativa para isso. O desenvolvimento impressionante da informática, juntamente com a proliferação generalizada em todos os domínios da computação, e o “nascimento” da Internet foram vitais para que isso acontecesse de facto. O evoluir destas tecnologias afectaram não só os particulares, mas também as empresas, permitindo o acesso via internet a um conjunto muito vasto de informações e a possibilidade de automatizar e processar algumas tarefas até então desempenhadas manualmente. No caso das empresas, o constante desenvolvimento de aplicações informáticas, para o processamento e automatização de tarefas, deu origem ao que agora são denominados de Sistemas Transaccionais. O desenvolvimento deste tipo de sistemas teve também alguma influência no *modus operandi* das organizações, atingindo um nível de importância tal que a estratégia de negócio de uma empresa é definida não só pelo evoluir do negócio e do mercado, mas também pelo evoluir dos seus Sistemas Transaccionais. No entanto, apesar destes sistemas conterem cada vez mais informação sobre o dia-a-dia da organização, esta poderá estar dispersa por diferentes aplicações, e provavelmente num estado inconsistente, reflectindo assim uma imagem que não será certamente a realidade da organização. Os Sistemas de *Data Warehousing* surgiram como mais um componente para um Sistema de Informação Empresarial com o objectivo de providenciar uma fonte de dados históricos validados, integrados, “limpos” e não-voláteis, que representem fielmente o estado actual da organização. Estes sistemas foram optimizados para acessos de consulta e permitem aos seus utilizadores a elaboração de relatórios e extracção de conhecimento que servirão de suporte ao processo de tomada de decisões. O projecto, análise e desenvolvimento deste tipo de sistemas requer alguma atenção, em especial nas tarefas inerentes à gestão do projecto, análise de requisitos e definição do âmbito, e no desenvolvimento de ferramentas de extracção, transformação e integração de dados. Existem neste momento na comunidade de *data warehousing* duas principais abordagens, uma preconizada por William H.

Inmon, mais centrada nos dados, e outra por Ralph Kimball, mais centrada no projecto. Ambas as abordagens são bastante válidas, mas a metodologia de Ralph Kimball tem obtido uma maior aceitação por parte do mercado devido à problemática existente na gestão de projectos de *data warehousing*. A percentagem de projectos de *data warehousing* que são mal sucedidos é tão elevada que se tem dada grande atenção à questão da gestão do próprio projecto. Assim, não é de surpreender que a metodologia de Ralph Kimball tenha ganho algum relevo nas organizações envolvidas no desenvolvimento de Sistemas de *Data Warehousing*. Em termos gerais, esta dissertação foi desenvolvida com base num estudo detalhado sobre o planeamento e a implementação de Sistemas de *Data Warehousing* em organizações reais. Abordaram-se, com particular ênfase, os principais aspectos relacionados com o planeamento, projecto, gestão e integração deste tipo de sistemas em meios empresariais. Complementarmente, e como forma de comprovar a aplicação no mundo real das técnicas de implementação estudadas, realizou-se uma avaliação das práticas correntes seguidas por algumas empresas de relevo no panorama nacional, e empregues no desenvolvimento de projectos de *data warehousing*.

Palavras-Chave: Bases de Dados, Gestão Projectos, Sistemas de Apoio à Decisão e Sistemas de *Data Warehousing*

Abstract

In the last years it is common to hear say that we are in the Age of Information. The widespread use of Information Technologies has contributed significantly to this. The remarkable development of Computer Science, along with the generalized proliferation in all computation domains, and the "birth" of the Internet were fundamental to this achievement. The evolution of these technologies affected significantly current activities of common people and companies, allowing for the access via Internet to large repositories of information, and the possibility to automate activities that were previously done manually. The constant development of software applications to process and automate tasks led to what we know today as Transactional Systems. The development of these systems also had some impact on the organization *modus operandi*, influenced strongly common business strategies through their own functionalities and services, which were systematically improved during their lifetime. However, despite the fact that these systems gather and store, day after day, more and more information about the organization's activities, such information may be distributed by different applications, probably in an inconsistent state, reflecting an image that is not the reality of the organization. Data Warehouse Systems emerged as a part of an enterprise information system providing now a new set of computational services that contribute effectively to added value to conventional data processing tasks through their abilities to explore historical, validated, integrated, clean and non-volatile sources of data. These systems were specially optimized for query access and allow their users to make reports and gather subject oriented information that will support business decision making processes. The project, analysis and development of Data Warehousing Systems require some particular attention, especially in tasks such as project management, requisites analysis, scope definition, and in the development of data warehouse populating tools. In the Data Warehousing area there are, essentially, two main approaches, a data-centred one by William H. Inmon, and a project-oriented approach proposed by Ralph Kimball. Both of them are valid approaches. However, the Ralph Kimball's methodology has gained larger acceptance within the organizations due to its special orientation to the data

warehouse project management problems. The percentage of unsuccessful data warehousing projects has been so high that data warehousing experts have been focused a large part of their today's efforts to project management issues. Thus, it is not a surprise the selection of the Ralph Kimball's methodology by the organizations that are involved with the project and development of a Data Warehousing System. In general terms, this thesis was developed based on a detailed study that was done about Data Warehousing Systems planning and implementation in real world organizations. Many topics were approached concerning with project design, implementation, management and integration of this systems in enterprise environments. Furthermore, it was performed a particular study about common practices that some relevant companies follow in the national "arena" of data warehousing systems. This study was very useful to prove the application of the implementation techniques and methodologies studied in this thesis to real world data warehousing systems projects.

Keywords: Databases, Project Management, Decision Support Systems, and Data Warehousing Systems

Índice

Agradecimentos	vii
Resumo	ix
Abstract.....	xi
Introdução	1
1.1 Sistemas de Apoio à Decisão	1
1.2 Motivação e Objectivos	5
1.3 Organização do Relatório	6
Planeamento Estratégico de Sistemas de <i>Data Warehousing</i>.....	9
2.1 Estratégia	9
2.2 Planeamento Estratégico de Sistemas de Informação.....	11
2.2.1 Evolução do Uso das Tecnologias da Informação nas Organizações	11
2.2.2 Evolução do Planeamento nas Organizações.....	11
2.2.3 Evolução do Planeamento de Sistemas de Informação.....	12
2.2.4 Planeamento de Sistemas de Informação	17
2.2.5 Factores de Sucesso e Insucesso no Planeamento de Sistemas de Informação	18
2.3 Planeamento Estratégico de Sistemas de <i>Data Warehousing</i>	20
2.3.1 Estratégia e Sistemas de <i>Data Warehousing</i>	20
2.3.2 Planeamento de Sistemas de <i>Data Warehousing</i>	23
Ciclo de Vida de um Projecto de um Sistema de <i>Data Warehousing</i>.....	25
3.1 Impacte de um SDW no Seio de uma Organização	25
3.1.1 Informação.....	25
3.1.2 Sistemas de Apoio à Decisão.....	27

3.1.3	Sistemas de <i>Data Warehousing</i>	29
3.2	Sucesso e Insucesso de um SDW	30
3.3	O Projecto de um SDW	35
3.3.1	Análise	36
3.3.2	Planeamento	37
3.3.3	Desenho.....	39
3.3.4	Desenvolvimento	39
3.3.5	Instalação e Formação	40
3.4	O Dia Seguinte ao Arranque do SDW.....	41
	Kimball, Inmon e Outros	43
4.1	A Abordagem de William H. Inmon.....	43
4.1.1	Projecto	43
4.1.2	Gestão	45
4.1.3	Implementação.....	46
4.2	A Perspectiva de Ralph Kimball	50
4.2.1	Um Plano para o Projecto	50
4.2.2	Gestão	56
4.2.3	Implementação.....	59
4.3	Outras Perspectivas	62
4.3.1	Sid Adelman	62
4.3.2	Michael Gonzales	65
	Gestão de Projectos de <i>Data Warehousing</i>.....	69
5.1	A Gestão de Projectos	69
5.1.1	Projecto	70
5.1.2	<i>Framework</i> Aplicada à Gestão de Projectos.....	71
5.1.3	<i>Framework</i> Aplicada à Gestão de Projectos – A Visão Europeia	78
5.2	Gestão de Projecto de <i>Data Warehousing</i>	81
5.3	Kimball e PMI	86
	Da Teoria à Prática.....	91
6.1	Casos de Estudo	91

6.2	Planeamento do Projecto.....	94
6.2.1	Definição do Projecto	94
6.2.2	Calendarização do Projecto	96
6.2.3	Orçamentação do Projecto.....	97
6.2.4	Constituição da Equipa de <i>Data Warehousing</i>	98
6.3	Gestão do Projecto	102
6.3.1	Monitorização do Projecto	102
6.3.2	Gestão do Âmbito	104
6.3.3	Plano de Comunicação.....	106
6.3.4	Qualidade.....	107
6.3.5	Factores Críticos	109
6.3.6	Dificuldades.....	110
6.4	Recolha e Análise de Requisitos	112
6.4.1	Entrevistas e <i>Workshops</i>	112
6.4.2	Sucesso ou Insucesso.....	112
6.5	Implementação	113
6.5.1	Linha de Acção Tecnológica	113
6.5.2	Linha de Acção de Dados.....	115
6.5.3	Linha de Acção das Aplicações	118
6.6	Instalação e Entrega do Projecto.....	119
6.7	Manutenção e Crescimento.....	120
6.8	Aspectos a Salientar.....	121
6.9	Resumo	122
	Conclusões	125
7.1	Planeamento Estratégico	125
7.2	Ciclo de Vida de um Projecto de um Sistema de <i>Data Warehousing</i>	127
7.3	Estudo e Análise de Diferentes Metodologias.....	129
7.4	Gestão do Projecto	131
7.5	Apreciação das Abordagens das Empresas	133
7.6	Apontamento Final.....	135
	Bibliografia.....	137

Referências WWW	151
------------------------------	------------

Índice de Figuras

Figura 1 - Evolução das abordagens do planeamento de SI [Lee et al. '03]	14
Figura 2 - Dados, Informação, Conhecimento, Acção	26
Figura 3 - Modelo de Desenvolvimento de <i>Software</i> (adaptado de [Boehm '88]).....	35
Figura 4 - METH2 [Inmon '02]	47
Figura 5 - METH3 / METH4 [Inmon '02]	49
Figura 6 - Ciclo de Vida Dimensional do Negócio [Kimball et al. '98]	50
Figura 7 - Grupos de Processos [PMI '00]	71
Figura 8 - Áreas do Conhecimento [PMI '00]	72
Figura 9 - Processo de Gestão de Projectos [APM '00a]	79
Figura 10 - Interação entre Fases [Duncan '95]	90
Figura 11 - Áreas Alvo	134

Índice de Tabelas

Tabela 1 - Níveis de Integração [Gottschalk et al. '01]	16
Tabela 2 – Resumo das tarefas a executar na elaboração do projecto segundo Kimball.....	56
Tabela 3 – Resumo das tarefas a executar na gestão de um projecto segundo Kimball	58
Tabela 4 – Funções típicas existentes num projecto de DW	67
Tabela 5 - Mapeamento entre Áreas do Conhecimento e Grupos de Processos [PMI '00].....	78
Tabela 6 - Empresas entrevistadas	92
Tabela 7 - Pontos Fortes dos Projectos SDW das Empresas Entrevistadas.....	123
Tabela 8 - Tópicos Importantes na Elaboração do Projecto, na Gestão e na Implementação de SDW	131

Capítulo 1

Introdução

1.1 Sistemas de Apoio à Decisão

Actualmente, assistimos por parte das empresas a uma procura desenfreada de novas soluções que lhes permitam melhorar os seus serviços e a sua forma de estar perante os mercados em que desenvolvem prioritariamente as suas actividades. As empresas tentam abranger, dia após dia, novos mercados que lhes garantam, basicamente, a sua “sobrevivência” e um crescimento sustentado ao longo dos anos. Paralelamente ao crescimento das suas actividades, verifica-se também, na generalidade dos casos, um aumento significativo do volume de informação gerada, tratada e armazenada. Porém, na generalidade dos casos, quanto maior for o volume de dados a tratar maior é o tempo de tratamento e de resposta às solicitações mais regulares. Sendo o tempo de resposta um factor importante no meio empresarial, uma das soluções adoptadas pelas empresas para o minimizar foi (e continua a ser) a informatização e a automatização dos processos existentes.

As *Tecnologias da Informação* (TI) vieram revolucionar a forma como as organizações funcionam, começando pela automatização de funções, facilitando operações como o registo de dados, emissão de documentos e controlo de máquinas [Porter et al. '85]. A evolução das TI foi tal forma contundente que potenciou a descoberta de novos métodos de trabalho, novas formas de comunicação, reduzindo o custo de algumas actividades e permitindo a descoberta de novas áreas de negócio.

No entanto, nem sempre o aparecimento de novas tecnologias tem sido aproveitado em todo o seu potencial. O uso adequado das novas tecnologias concede às organizações vantagens sobre as suas concorrentes, quer através da optimização de funções e tarefas, que permitirão obter uma redução do custo do produto ou serviço ao cliente, quer através da angariação de dados sobre o cliente permitindo um melhor conhecimento das suas necessidades e, como tal, criando a oportunidade de melhor direccionar os esforços da organização com vista à sua satisfação. Com o intuito de atingir estes objectivos, as organizações têm investido fortemente no desenvolvimento de aplicações específicas ao seu negócio. No entanto, o desenvolvimento destas aplicações nem sempre são alvo de um correcto planeamento. A ausência de planeamento tem sido a principal razão para a elevada taxa de insucesso no desenvolvimento de tais aplicações, normalmente designadas de sistemas transaccionais ou sistemas de informação [Min et al. '99], colocando em causa o planeamento estratégico do negócio, uma vez que dificilmente estes sistemas atingirão os objectivos pretendidos.

A proliferação de aplicações informáticas ou de sistemas transaccionais tem originado um contínuo acumular de dados, o que não revela necessariamente um aumento de conhecimento sobre o negócio. Uma das razões principais prende-se com o facto destas aplicações terem sido desenvolvidas autonomamente, não tendo sido construído um modelo de dados global da empresa, resultando assim no aparecimento de redundância e, conseqüente, inconsistência de informação. A inconsistência de informação é a principal causa para o não aproveitamento das potencialidades das novas tecnologias da informação e um potencial problema para o negócio, dado que as decisões assentaram sobre informação incongruente. Outra razão é o facto de os sistemas transaccionais, por definição, servirem de suporte às organizações, registando e processando as operações do dia-a-dia, tornando-os inadequados para a realização de tarefas de

análise dos dados, correlação de dados ou até da elaboração de relatórios que reportem a evolução da organização entre períodos de tempo diferentes [Gupta '97].

O surgimento de *Sistemas de Apoio à Decisão* (SAD) veio colmatar algumas falhas existentes nos sistemas transaccionais relativamente à exploração dos dados existentes. Os SAD são desenvolvidos em plataformas orientadas à consulta, e têm como principais objectivos [Shim et al. '02]:

- Ajudar, aconselhando, no processo de tomada de decisões.
- Disponibilizar um repositório de dados para consulta.
- Disponibilizar um interface amigável.
- Facilitar o desenvolvimento de *queries*.

No entanto, o desenvolvimento dos SAD dentro das organizações não seguiram um plano estruturado de integração de dados, o que fez com que muitos destes sistemas se baseassem em apenas numa parte dos dados existentes na organização, verificando-se uma vez mais a existência de diferentes “verdades” sobre o mesmo assunto. Tal incongruência entre relatórios ou análises efectuadas em diferentes SAD apenas revelou que estes sistemas não serviam como suporte às tomadas de decisão dos gestores de uma organização.

Uma das formas para solucionar a integração dos dados de diferentes sistemas ou aplicações é o desenvolvimento e implementação de *Sistemas de Data Warehousing* (SDW) numa organização. No entanto, à imagem do aparecimento dos primeiros *Sistemas de Informação* (SI), os SDW devem ser alvo de um planeamento cuidado, de forma a garantir o sucesso do desenvolvimento de um sistema de elevada complexidade e de vital importância para o futuro da organização, uma vez que este sistema disponibilizará um repositório de dados históricos integrados, validados e “limpos”. Este repositório vai servir de base de consulta e de extracção de conhecimento que ajudará no processo de tomada de decisões que afectarão a sobrevivência e o desenvolvimento da organização.

Os SDW surgiram assim como um “complemento” natural aos sistemas operacionais das empresas, uma vez que estes estão orientados à análise de dados através da disponibilização de um repositório integrado de dados históricos, ao invés dos sistemas operacionais que estão orientados ao processamento de transacções e ao suporte da actividade do negócio. E como tal, permitem que grandes volumes de informação, provenientes de diferentes fontes, sejam extraídos, tratados, e armazenados num só repositório de dados possibilitando desta forma o desenvolvimento de processos de análise mais rápidos e efectivos.

Os SDW destinam-se, essencialmente, a utilizadores com capacidade de decisão sobre o futuro da empresa que, com ajuda de ferramentas especializadas, conseguem analisar padrões e variações nos dados, extraíndo daí os elementos necessários para responder, em tempo útil, às mais diversas questões e desenvolver as bases para o suporte quotidiano dos seus processos de tomada de decisão. Uma vez que estes sistemas albergam dados provenientes de diversas fontes de informação empresariais, é possível também correlacionar os mais diversos tipos de dados que até então, somente com os sistemas operacionais, eram um pouco difíceis de correlacionar no sistema de informação [Chaudhuri et al. '96, Hammer et al. '95].

Dadas as características dos SDW, as organizações começaram a pensar na adopção deste tipo de sistema como um investimento no futuro. Estes sistemas têm como principais vantagens :

- o potencial de servir como sistemas de apoio à decisão;
- criação de um sistema de armazenamento de dados tratados;
- servir de base para ferramentas de análise de dados.

No entanto, o desenvolvimento de SDW não pode ser abordado como se de um projecto normal de *software* se tratasse, tal atitude tem originado uma alta percentagem de insucesso no desenvolvimento de SDW. É necessário reconhecer que um projecto de um SDW é um projecto que necessita de um planeamento elaborado, de respeitar um conjunto de regras específicas e de ter especial atenção aos factores críticos deste tipo de projectos que, normalmente, elevam o grau de risco do projecto.

1.2 Motivação e Objectivos

A evolução dos Sistemas de Informação nas ultimas décadas originou um crescente acumular de informação em bases de dados empresariais, possivelmente instaladas em plataformas heterogéneas, que com o decorrer do tempo tornam o seu manuseamento, a sua análise e garantia de inconsistência um problema de proporções preocupantes [Chaudhuri et al. '96].

Sendo assim, tornou-se necessário o desenvolvimento de um repositório de dados que estivesse optimizado para processos de interrogação mais efectivos, independente dos sistemas operacionais, mas que destes obtivesse a sua informação. Esta informação é extraída, tratada e guardada integradamente no que é chamado de *Data Warehouse* (DW). No entanto, o projecto de desenvolvimento de um sistema que acede aos variados sistemas operacionais, extrai os dados aí guardados, transforma-os para serem inseridos numa base de dados que representa o modelo de dados global da organização é um projecto complexo e moroso.

Com base nos pressupostos referidos anteriormente, definiu-se como objectivos principais para esta dissertação os seguintes pontos:

- Analisar a evolução dos Sistemas de Informação e respectivo planeamento.
- Analisar o ciclo de vida e o impacte de um projecto de desenvolvimento de um SDW numa organização.
- Analisar as diferentes metodologias existentes de desenvolvimento de SDW.
- Estudar a problemática inerente à Gestão de Projectos de SDW.
- Efectuar um estudo crítico das práticas correntes, no mercado nacional, no desenvolvimento destes Sistemas.

Para tal ser possível, foi necessário estudar a evolução dos Sistemas de Informação Empresariais, identificando as diferentes fases dessa evolução, quais os motivos que a originaram juntamente com a descoberta da necessidade de um planeamento estratégico associado ao desenvolvimento dos SI, e em particular de um SDW.

Uma vez definido o projecto de desenvolvimento do SDW, foi feita uma análise crítica às propostas existentes de autores conceituados na área de *Data Warehousing*, com particular incidência nas abordagens propostas por William H. Inmon e Ralph Kimball. Foram também entrevistadas diversas empresas envolvidas em projectos de *data warehousing* com o intuito de perceber quais as metodologias seguidas para a gestão e desenvolvimento do projecto, e até que ponto essas metodologias eram ou não semelhantes às estudadas.

1.3 Organização do Relatório

Além do presente capítulo, esta dissertação integra um conjunto de mais seis capítulos, organizados da seguinte forma:

- **Capítulo 2**

Planeamento Estratégico de Sistemas de *Data Warehousing*

Este capítulo tem como principal objectivo analisar a evolução dos SI e o planeamento associado ao seu desenvolvimento, prestando especial atenção aos SDW.

- **Capítulo 3**

Ciclo de Vida de um Projecto de um Sistema de *Data Warehousing*

Uma vez definido um SDW, foi estudado e avaliado, no decorrer deste capítulo, o impacto que um projecto desta natureza tem nas organizações onde são implementados e definidas as fases e os intervenientes do projecto.

- **Capítulo 4**

Kimball, Inmon e Outros

Este capítulo dedica-se à análise de duas das mais importantes metodologias no desenvolvimento de SDW, avaliando pontos fortes e fracos de cada metodologia e complementado os pontos fracos com perspectivas de outros autores.

- **Capítulo 5**

Gestão de Projectos de *Data Warehousing*

Dada a elevada taxa de insucesso no desenvolvimento de projectos de *data warehousing*, este capítulo foi dedicado ao factor que maior influência tem sobre o resultado final do projecto, ou seja, a Gestão de um Projecto de *Data Warehousing*. Foi apresentada a *framework* da *Project Management Institute* (PMI) aplicada à gestão de projectos e feita a comparação com a metodologia de Kimball.

- **Capítulo 6**

Da Teoria à Prática

Após a realização de diversas entrevistas a gestores de projecto de *data warehousing*, são expostas, neste capítulo, as metodologias seguidas por diversas empresas que estão ou estiveram inseridas em projectos de *data warehousing*. As metodologias internas são então comparadas com a metodologia de Kimball, sendo assim analisada a ponte entre a teoria e a prática mais comum na área de *Data Warehousing*.

- **Capítulo 7**

Conclusões

Neste capítulo faz-se uma análise crítica ao desenvolvimento de projectos de *data warehousing*, especialmente quando inseridos numa estratégia global de desenvolvimento dos SI. São também analisadas resumidamente as diferentes perspectivas dos diversos autores estudados assim como a aplicabilidade prática das metodologias que advogam. A finalizar, é feito um breve resumo da prática comum na área de *Data Warehousing* no mercado nacional.

Capítulo 2

Planeamento Estratégico de Sistemas de *Data Warehousing*

2.1 Estratégia

A utilização de um conjunto de meios e de planos para atingir um fim é normalmente denominada de Estratégia. Esta, é utilizada nas mais diversas situações e ambientes por indivíduos, grupos, ou empresas.

A utilização de uma estratégia para atingir um determinado fim apenas tem sucesso se esta tiver sido planeada, daí que, uma vez definido um conjunto de objectivos a atingir num determinado prazo, torna-se essencial delinear uma estratégia que nos providencie uma direcção a seguir, conseguindo assim uma concentração de esforços na obtenção desses objectivos e exigindo perseverança e flexibilidade [Boar '97].

A eterna luta das empresas é a tentativa de obtenção de alguma vantagem sobre as suas mais directas concorrentes. Como tal, o planeamento estratégico do negócio visa a obtenção e manutenção de vantagens que permitirão obter uma maior satisfação dos clientes, criando assim um maior distanciamento dos concorrentes. O planeamento estratégico deve também tentar eliminar ou minimizar as vantagens obtidas pelos concorrentes de forma a manter e atrair mais clientes [Boar '97].

Existem cinco tipo de vantagens que se poderão obter através de um planeamento estratégico: custo, valor acrescentado, focus, rapidez, flexibilidade.

As vantagens do tipo Custo estão normalmente associadas à comercialização de um produto ou serviço a um valor inferior ao da concorrência, contribuindo assim de forma significativa para a obtenção de uma vantagem sobre as demais empresas.

Através da disponibilização de um produto ou serviço com características ou qualidades acima da média, as organizações obtêm a vantagem de disponibilizar um produto ou serviço com valor acrescentado por um valor semelhante ao de um produto normal. A orientação das empresas para a satisfação das necessidades e exigências do cliente é determinante para a manutenção dos mesmos e também para captação de novos clientes, especialmente nos mercados onde os clientes são cada vez mais exigentes. Outra vantagem, essencial para fazer frente à concorrência cada vez mais feroz do mercado em que a organização se insere, é a rapidez e prontidão com que são satisfeitas as necessidades de um cliente.

A flexibilidade com que uma empresa se adapta às constantes mudanças do dia-a-dia, às mudanças das necessidades dos clientes e à constante evolução dos produtos é fundamental para a manutenção das quatro vantagens atrás descritas. Esta última vantagem é a única que nenhuma organização pode conquistar a outra [Boar '97].

2.2 Planeamento Estratégico de Sistemas de Informação

2.2.1 Evolução do Uso das Tecnologias da Informação nas Organizações

A utilização das TI nas organizações tem evoluído, satisfazendo cada vez mais requisitos da área do negócio. Essa evolução tem sido classificada segundo três eras : processamento de dados, SI de apoio à gestão e SI estratégicos [Ruohonen '96].

A era do processamento de dados teve início na década de 60 com o aparecimento dos primeiros computadores. A utilização dos computadores permitiu obter ganhos de produtividade uma vez que estes vieram automatizar processos e processar informações de forma mais célere e eficiente.

O desenvolvimento de SI que suportam e integram diferentes áreas da organização registando e processando dados, dá origem à segunda era, com início na década de 70. Esta era está assente no propósito de armazenamento de informação para consulta e análise.

A terceira era tem por base a crescente importância de um SI numa organização, atingindo este o estatuto de SI estratégico uma vez que para além de suportar todas as operações da organização serve também de suporte e de fonte de informação aos decisores da organização. Esta era começa na década de 80 e persiste até aos dias de hoje.

De notar que as propriedades de uma era não são exclusivas da mesma, bem pelo contrário. As propriedades das eras passadas vão permanecendo e vão sendo alargadas ou desenvolvidas dando origem a novas etapas de desenvolvimento [Peppard et al. '04].

2.2.2 Evolução do Planeamento nas Organizações

Para ser possível efectuar uma análise sobre a evolução do planeamento dos SI é primeiro necessário analisar a evolução do planeamento do negócio nas organizações, uma vez que o

planeamento estratégico de uma empresa é a base de trabalho para a definição dos requisitos que os SI deverão suportar.

O planeamento de negócio envolve cinco fases, Planeamento Orçamental e Financeiro, Planeamento a Longo Prazo, Planeamento Estratégico de Negócio, Planeamento Estratégico da Organização, Gestão Estratégica [Ruohonen '96].

A primeira fase, referente ao Planeamento Orçamental e Financeiro, está mais preocupada com a gestão eficiente dos recursos disponíveis e com o cumprimento do orçamento atribuído. No Planeamento a Longo Prazo, segunda fase, são utilizados os dados acumulados para melhor compreender o passado e como tal planear o futuro num prazo mais estendido.

Estas duas primeiras fases assentam bem em negócios em que a evolução é limitada e em que a concorrência é estável. No entanto, a incerteza e a evolução são cada vez mais o dia-a-dia das organizações, daí que tenham surgido novos modelos de planeamento. O Planeamento Estratégico do Negócio e da Organização são consequência da crescente incerteza e competitividade do ambiente empresarial. O Planeamento Estratégico de Negócio analisa uma determinada área de negócio, enquanto que o Planeamento Estratégico da Organização visa definir regras e linhas de acção que providenciem sinergias entre as diferentes áreas de negócio de uma organização.

À medida que surgem tempos cada vez mais incertos, de extrema concorrência, competitividade e inovação, o planeamento do negócio ganha ainda mais importância, estando em permanente revisão e comparação com a restante concorrência, em tópicos como estrutura, cultura organizacional e recursos humanos. A Gestão Estratégica visa a investigação e a criação de cenários como forma de preparação para um futuro incerto.

2.2.3 Evolução do Planeamento de Sistemas de Informação

A par da evolução do planeamento do negócio nas organizações, a utilização das TI, como ferramentas de suporte ao alcance dos objectivos propostos, tem também evoluído. A cada vez maior importância das TI numa organização tem como principal motivo a constante evolução do

mercado, a automatização de tarefas e a automatização de processamento de dados que estas providenciam, e as possíveis vantagens competitivas que o desenvolvimento de SI podem originar.

No entanto, este último motivo é discutível. O desenvolvimento de SI trará vantagens competitivas, mas estas podem facilmente ser copiados pelos concorrentes, o que fará com que sejam anuladas as vantagens alcançadas pelo seu desenvolvimento dentro de uma organização. Tal é possível devido ao facto de não existirem patentes para os SI desenvolvidos e também pela impossibilidade de se manter em segredo um SI inovador, especialmente quando este é utilizado por clientes ou fornecedores [Peppard et al. '04].

Chega-se então à conclusão que o desenvolvimento de SI apenas proporcionará vantagens competitivas se estas conseguirem ser alcançadas de forma sustentada e constante, nomeadamente através do desenvolvimento constante dos SI ou através da inclusão de novas funcionalidades requeridas pelos utilizadores ou clientes. O desenvolvimento dos SI deve por isso obedecer a um planeamento cuidadoso, de forma a atingir os objectivos propostos e desejados. O planeamento estratégico de SI é o processo de identificação de um conjunto de aplicações informáticas que suportarão as actividades da organização permitindo que esta alcance os seus objectivos [Lederer et al. '88].

O planeamento de SI também teve a sua evolução (Figura 1) no início, o planeamento era orientado à criação de um conjunto de aplicações que aumentariam o desempenho e a eficiência de algumas tarefas através da automatização das tarefas e do processamento de dados. O problema associado a este tipo de planeamento era a ausência de uma visão de longo prazo e a falta de integração entre as diversas aplicações desenvolvidas. Era comum uma empresa proceder à informatização dos sectores por fases, desenvolvendo aplicações que respondessem aos requisitos específicos desse sector, não havendo uma visão global das necessidades da organização. A consequência deste tipo de abordagem é a proliferação de aplicações autónomas, como gestão de encomendas, gestão de stocks, gestão de facturação, e gestão contabilística que embora autónomas armazenam informação comum, como por exemplo dados sobre clientes, produtos ou matérias primas, havendo assim redundância de informação e o risco de esta informação se tornar inconsistente.

Para colmatar a falha da abordagem anterior o planeamento de SI evoluiu, começando por ter uma base de trabalho que são os requisitos de negócio e os objectivos a alcançar, determinados pelo planeamento do negócio. Os SI são, assim, desenvolvidos como uma ferramenta que ajudará à obtenção dos objectivos previamente definidos no plano estratégico da organização.

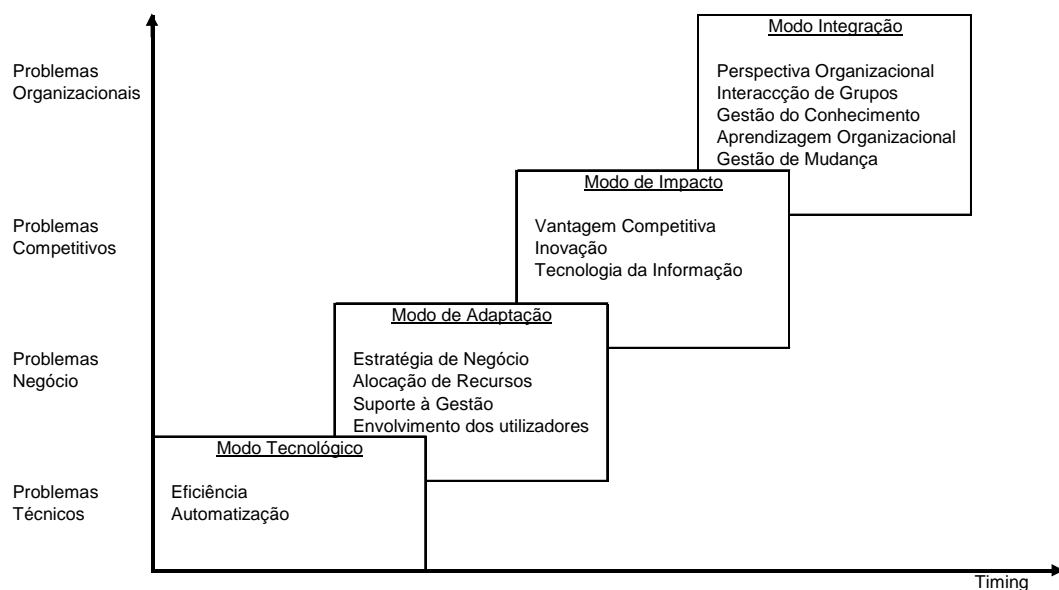


Figura 1 - Evolução das abordagens do planeamento de SI [Lee et al. '03]

Esta abordagem, inserida nas metodologias orientadas aos modelos de dados, foi classificada como reactiva e demasiado morosa, uma vez que exige a análise detalhada de todos requisitos de negócio, comprometendo assim o desenvolvimento dos SI. Uma nova abordagem surgiu na qual são previamente identificadas as aplicações a desenvolver que poderão garantir à organização a obtenção de alguma vantagem competitiva frente às suas concorrentes. A obtenção de tais vantagens poderá influenciar a estratégia de negócio. Como tal, o planeamento de SI deixa de ser uma consequência do planeamento do negócio para passar a ter uma relação de bidireccionalidade, isto é, o desenvolvimento de SI poderá originar algum impacte ou até, ser de

importância crucial para o futuro da organização e como tal alterar a sua estratégia de desenvolvimento [Kearns et al. '00].

A abordagem mais recente tem em linha de conta o facto do desenvolvimento de SI originar impacte na organização, através da alteração dos procedimentos, tarefas, aptidões ou até na sua estrutura e na sua cultura. Daí que, para se conseguir, com sucesso, desenvolver os SI adequados à organização é preciso primeiro estudar a estrutura e cultura organizacional a fim de adaptar os SI à organização e como tal atingirem-se os objectivos da organização mais facilmente.

É assim possível concluir que o planeamento de SI tem-se adaptado à evolução natural das organizações, seguindo as cinco fases enunciadas anteriormente sobre a evolução do planeamento do negócio. A integração entre os SI e o negócio tem evoluído da mesma forma e tal evolução repercute-se na organização.

Resumindo, existem dois tipos de abordagens diferentes para o planeamento de SI. Uma é baseada na premissa que é possível planear o desenvolvimento estratégico de SI e outra é baseada na premissa que os SI evoluem com o tempo e com a evolução da própria organização. Assim, estes não podem ser objecto de um planeamento a longo prazo, uma vez que existem mudanças a nível tecnológico, do negócio ou da gestão, que possivelmente afectarão qualquer planeamento a longo prazo [Peters et al. '02].

Mecanismos de Integração	Integração Administrativa	Integração Sequencial	Integração Recíproca	Integração Total
Objectivo da Integração	Administrativa e não estratégica	Suporte à Estratégia de Negócio	Suporte e influência à Estratégia de Negócio	Desenvolvimento conjunto das Estratégias de Negócio e dos SI
Funções dos SI	Técnicas e não estratégicas	Recurso de suporte à estratégia de negócio	Recurso de suporte e influência à estratégia de negócio	Critico para a sobrevivência a Longo Prazo da Organização
Funções do Responsável pelos SI	Administrador funcional e responsável pelo suporte	Especialista que planeia os SI para implementar a estratégia de negócio	Especialista que contribui para o planeamento e desenvolvimento da estratégia de negócio	Membro da Administração da Organização
CrITÉrios de Performance	Eficiência e redução de custos	Contribui para a implementação da estratégia de negócio	Qualidade das informações que influenciam a estratégia de negócio	Impacte de Longo Prazo na organização
Razões para o desenvolvimento dos SI	Necessidade de automatização dos processos administrativos	Obtenção dos objectivos do negócio	Objectivos do negócio e dos SI considerados juntamente	Os SI são críticos para o sucesso da estratégia de negócio
Participação da Administração no Planeamento dos SI	Raramente	Ocasionalmente	Frequentemente	Quase Sempre
Participação dos utilizadores no Planeamento dos SI	Raramente	Ocasionalmente	Frequentemente	Quase Sempre
Participação do Responsável dos SI no planeamento do Negócio	Raramente	Ocasionalmente	Frequentemente	Quase Sempre
Avaliação de novas Tecnologias	Raramente	Ocasionalmente	Frequentemente	Quase Sempre
Status do Responsável pelos SI (níveis abaixo do Administrador)	Quatro ou mais	Três	Dois	Um

Tabela 1 - Níveis de Integração [Gottschalk et al. '01]

Embora existam estes dois pontos de vista, há também quem proponha um compromisso entre os dois, defendendo que o planeamento dos SI deva existir, mas que este deve ser efectuado de forma informal e incremental no sentido de garantir flexibilidade, criatividade e pensamento estratégico com o objectivo de incorporar quer estratégias emergentes quer estratégias planeadas [Salmela et al. '02].

2.2.4 Planeamento de Sistemas de Informação

O planeamento estratégico de SI está dividido em fases que por sua vez estão divididos em estágios e actividades [Mentzas '97]:

- Planeamento do processo de planeamento dos SI (*strategic awareness*), fase em que o principal objectivo é consciencializar os intervenientes na necessidade de elaborar uma estratégia de planeamento para os SI, e onde as principais tarefas a executar são:
 - Determinação dos tópicos essenciais ao planeamento.
 - Definição dos Objectivos do Planeamento.
 - Organização das equipas de Planeamento.
 - Obtenção do Apoio e Empenho da Administração.
- Análise do ambiente (*situation analysis*). Nesta fase, os objectivos essenciais são a análise dos diversos factores que afectam a organização, quer a nível interno quer externo, no âmbito dos SI.
 - Análise dos sistemas de negócio actuais.
 - Análise dos sistemas organizacionais actuais.
 - Análise dos sistemas de informação actuais.
 - Análise do ambiente externo de negócio.
 - Análise do ambiente externo das TI.
- Concepção de Alternativas Estratégicas (*strategy conception*), fase onde são desenvolvidas possíveis estratégias a adoptar pela organização, que confirmam oportunidades de obtenção de vantagens competitivas.
 - Identificação dos principais objectivos para as TI.
 - Identificação de oportunidades de melhoramento .
 - Avaliação das oportunidades de melhoramento.
 - Identificação das estratégias de alto-nível para as TI.

- Selecção da Estratégia (*strategy formulation*) visa a análise das diversas estratégias formuladas no ponto anterior tendo em conta a estrutura actual e a estrutura necessária para os implementar.
 - Identificação de novos processo de negócio.
 - Identificação de novas arquitecturas de TI.
 - Identificação de novos projectos.
 - Identificação das prioridades para os novos projectos.
- Planeamento da Implementação Estratégica (*strategy implementation planning*) é a fase onde são definidas e orçamentadas as acções a desenvolver, o seu calendário e os recursos necessários para a sua execução.
 - Definição da abordagem para a gestão de mudanças.
 - Definição do plano de acção.
 - Avaliação do plano de acção.
 - Definição dos procedimentos de acompanhamento e controlo.

2.2.5 Factores de Sucesso e Insucesso no Planeamento de Sistemas de Informação

O planeamento e desenvolvimento de SI é um processo sobre o qual incidem alguns factores que poderão fazer com que o projecto se torne um sucesso ou um insucesso. Vários estudos reportam taxas de insucesso no desenvolvimento de SI na ordem de dois terços a três quartos [Adelakun '96]. O desenvolvimento de um SI é considerado um fracasso quando este não é concluído, quando não atinge os objectivos propostos, quando não atinge as expectativas dos *stakeholders*¹, ou quando o custo com o desenvolvimento do SI ultrapassa os possíveis benefícios que este permitirá alcançar [Flowers '96].

¹*stakeholders* – partes interessadas no projecto. Ex: patrocinadores, clientes, utilizadores, gestores, membros da equipa, entre outros.

Antes de se analisar quais os factores que influenciam o resultado do desenvolvimento dos SI é preciso identificar que tipos de falhas existem para melhor determinar as suas causas. Basicamente estão identificados quatro tipos de falhas [Lyytinen et al. '87] :

- Falha de Correlação – Quando não são atingidos os objectivos definidos previamente no planeamento dos SI.
- Falha no Processo de Desenvolvimento – Quando os SI não são desenvolvidos no orçamento e calendário previsto. Este tipo de falha está normalmente associado à ineficiente gestão do projecto.
- Falha de Interação – Quando os utilizadores não estão satisfeitos com os SI desenvolvidos. Esta insatisfação não implica necessariamente a não utilização dos SI. Os SI podem ter uma taxa de utilização elevada e os utilizadores continuarem insatisfeitos. Tal acontece quando não existe outra possibilidade a não ser a utilização dos SI existentes.
- Falha de Expectativa – Este tipo de fracasso acontece quando os SI desenvolvidos não correspondem às expectativas dos *stakeholders*.

Alguns dos factores críticos que, a nível organizacional, têm influência sobre o planeamento e desenvolvimento de SI são [Yeo '02]:

- Cultura organizacional hostil, pode acontecer quando o projecto visa alterar o funcionamento de uma organização sem que os seus funcionários tenham sido incorporados no projecto, demonstrando assim resistência às possíveis alterações.
- Pressões políticas. Os elementos que estão no desenvolvimento do projecto poderão sofrer pressões provenientes de variados sectores da organização com vista a satisfação de alguma intenção.
- Influências. A equipa de desenvolvimento, ou os seus membros, poderão estar sujeitos a influências por parte de elementos internos ou externos à organização.
- Empenho da administração inadequado. Dada o risco inerente a este tipo de projectos, a falta de apoio e cometimento por parte da administração poderá influenciar negativamente os restantes membros da organização e como tal comprometer o projecto.

Os factores que podem influenciar a gestão do projecto, do planeamento e do desenvolvimento de SI são [Yeo '02] :

- Planeamento orientado à tecnologia.
- Ênfase na tecnologia ao invés das relações humanas.
- Subavaliação da complexidade do projecto.
- Gestão inadequada dos *stakeholders*.
- Ausência de consultores especializados.
- Adopção de uma solução tecnológica para um problema de gestão.
- Fraca competência para a gestão do projecto e da equipa.
- Selecções inadequadas (recursos humanos e tecnológicos).

É necessário um cuidado especial no planeamento dos SI, onde devem ser tomadas as precauções adequadas (acções preventivas) para evitar a ocorrência de um destes factores e devem também ser definidas as acções correctivas no caso de uma ocorrência inesperada que possa comprometer o projecto. A principal consequência inerente ao planeamento inadequado de SI não é apenas financeiro, é também a possibilidade de não se ter aproveitado as oportunidades que entretanto surgiram [Min et al. '99].

2.3 Planeamento Estratégico de Sistemas de *Data Warehousing*

2.3.1 Estratégia e Sistemas de *Data Warehousing*

Basicamente, um SDW pode ser visto como um componente de um SI. É um sistema que integra dados provenientes de diferentes sistemas fontes com o intuito de fornecer um repositório de dados históricos validados, integrados e “limpos”, sobre o qual se possam efectuar consultas, elaborar relatórios e até alimentar outros sistemas que permitirão servir de base aos processos de tomada de decisão das organizações.

A crescente importância dos SDW deve-se em parte ao actual panorama dos SI. A consequência do planeamento dos SI, segundo os modos anteriormente analisados, é a proliferação de aplicações informáticas destinadas à resolução de um problema e que foram desenvolvidas sem um planeamento ajustado às necessidades da organização. A evolução do planeamento dos SI do modo tecnológico e de adaptação para o modo de impacte e integração veio originar o aparecimento dos SDW que, dadas as suas características, começaram a obter cada vez mais aceitação por parte das organizações. Estudos revelam que de um conjunto de 250 empresas contactadas em 1995, 95% tencionavam construir um SDW, quando, um ano antes, destas mesmas empresas, apenas 15% tencionava construir um SDW [ADS '95]. Tal pode-se justificar com o sucesso obtido pelas organizações pioneiras no desenvolvimento de SDW, influenciando deste modo as restantes organizações.

A proliferação de aplicações informáticas independentes originou redundância de informação, a consequente inconsistência, e diferente representação da mesma informação. A juntar a estes problemas, a inadequação dos sistemas transaccionais para suportar análises complexas de dados contribuíram para que as organizações decidissem avançar para o desenvolvimento de SDW [Subramanian et al. '96]. A utilização correcta dos SDW permite às organizações adquirir conhecimentos através do acesso aos dados históricos, adaptar estratégias com os dados do presente e orientar a empresa para o futuro através da previsão ou da detecção de padrões ou sazonalidades [Boar '97].

Embora o planeamento e desenvolvimento de um SDW seja muito custoso, o potencial retorno que podemos obter com a utilização do SDW é também elevado. Um SDW é um sistema em constante crescimento, quer a nível de informação armazenada, quer a nível de potenciais utilizadores, não devendo por isso ser considerado como um sistema que apenas gera despesas, mas sim como um sistema que serve de suporte às decisões tomadas no dia-a-dia pelos seus utilizadores. Sendo assim, as decisões tomadas passam a ter um fundamento válido e facilmente comprovado ao invés das decisões tomadas por palpite ou intuição. Devemos por isso considerar um SDW como um sistema de importância estratégica para uma organização. Um SDW tem o potencial de permitir às organizações atingir uma ou mais das vantagens enunciadas numa estratégia para alcançar vantagem competitiva sobre as organizações concorrentes.

A primeira vantagem enunciada que uma organização pode obter através do planeamento estratégico é referente ao custo do produto ou serviço disponibilizado. Um custo inferior conferirá à organização uma vantagem competitiva frente às suas concorrentes. Tal vantagem pode ser alcançada através da utilização da informação existente no SDW para suportar decisões como a escolha do melhor fornecedor tendo em conta o preço ou do prazo de entrega ou condições de pagamento.

A disponibilização de um produto ou serviço de melhor qualidade ou com uma mais valia para o cliente é também uma vantagem que pode ser obtida através da utilização do SDW. Uma vez que o SDW serve como repositório de dados não-voláteis integrados, é possível, por exemplo, através do uso de ferramentas de análise de dados estabelecer um perfil do cliente e sugerir ou aconselhar a compra de produtos ou serviços de interesse para o cliente. A organização está desta forma focada nas necessidades do cliente, tentando satisfazê-las e como tal fidelizando o cliente, ficando desta forma em vantagem comparativamente às suas concorrentes.

A rapidez e flexibilidade são também vantagens desejadas pelas organizações e que os SDW poderão ajudar a atingir. Os SDW estão otimizados para acessos de consulta e análise de dados, proporcionando assim um ambiente onde a recolha de informação e obtenção de conhecimento é célere, tornando a organização flexível ao ponto de ser possível redefinir estratégias de posicionamento no mercado num curto espaço de tempo, estando desta forma um passo à frente da concorrência e como tal obtendo o máximo partido das mudanças constantes que o mercado proporciona.

Resumindo, o desenvolvimento de um SDW vai permitir [Poe et al. '98]:

- Uma vista integrada e completa de toda a organização.
- Ter acesso a dados históricos da organização.
- Ter acesso a uma fonte de dados verosímeis dentro da organização.
- Facilitar os processos de tomada de decisão sem sobrecarregar os sistemas operacionais.

2.3.2 Planeamento de Sistemas de *Data Warehousing*

O planeamento de desenvolvimento de um SDW está normalmente inserido num contexto de maior dimensão. Isto é, o desenvolvimento de um SDW é consequência de um planeamento estratégico de SI que tem como finalidade suportar e ajudar uma organização a atingir os seus objectivos, objectivos esses previamente definidos no plano estratégico de negócio.

Considerando um SDW como um componente de um SI, o planeamento deste deve ser integrado no planeamento do SI. Daí que, as tarefas associadas ao planeamento do SDW são semelhantes, ou até decorrentes, das tarefas definidas no planeamento do SI.

No entanto, existem também alguns casos em que a disparidade entre SI e plataformas é tal, que o intercâmbio de informação ou dados é uma tarefa de extrema complexidade. Nestes casos, em que a reformulação dos SI existentes são tarefas de demasiada complexidade e grandeza, o conceito de *data warehousing* emerge, assim, como uma possível solução para o armazenamento integrado de dados provenientes desses SI dispares.

O município de St. Louis nos Estados Unidos da América é um caso destes, em que a estratégia definida para uma melhor integração dos dados existentes foi o desenvolvimento de um *data warehouse* que iterativamente fosse albergando os dados necessários à consulta dos mais diversos serviços e departamentos. Neste caso, foi delineado um planeamento estratégico para o desenvolvimento do SDW constituído pelos seguintes passos [Subramanian et al. '97]:

1. Constituição da equipa que elaborará o plano para o desenvolvimento do SDW constituída por elementos que identificarão pontos críticos e reunirão consenso sobre o plano a executar.
2. Sessões de esclarecimento a todos os intervenientes no processo como forma de os sensibilizar e incorporar no projecto e refinar objectivos do projecto.
3. Elaboração e análise de questionários sobre as actuais funções desempenhadas, as necessidades de informação de cada departamento, meios de comunicação inter-departamental e meios informáticos utilizados para a execução das tarefas diárias.

4. Elaboração e análise de questionários sobre as infra-estruturas tecnológicas existentes e possíveis melhoramentos.
5. Sessões de esclarecimento sobre as actuais funcionalidades e capacidades dos SI existentes como forma de educação dos elementos que participam na elaboração do plano estratégico.
6. Formação dos membros da equipa e dos gestores nos tópicos relativos a *data warehousing*.
7. Criação de uma matriz sobre os dados consultados nos diversos departamentos (incluindo o tipo e frequência do uso).
8. Criação de relatórios sobre iniciativas correntes e futuras de criação ou desenvolvimento de SI em cada departamento.
9. Criação de um relatório sobre potenciais SI que abranjam diversos departamentos, identificando procedimentos interdepartamentais e respectivas necessidades de informação.
10. Informação dos diversos departamentos sobre os dados recolhidos e que ajudará a confirmar a sua autenticidade para a elaboração do plano final.
11. Estudo das infra-estruturas tecnológicas existentes.
12. Avaliação e escolha de topologias alternativas para a futura infra-estrutura tecnológica.
13. Estudo das práticas adoptadas em organizações semelhantes, esta medida serve também para tentar convencer os mais cépticos quanto ao sucesso atingido por outras organizações em projectos semelhantes.
14. Revisão da documentação existente sobre os SI existentes e de iniciativas passadas mal sucedidas.
15. Análise das recomendações provenientes dos diversos departamentos quanto aos dados recolhidos no processo de levantamento de requisitos.
16. Elaboração do plano de desenvolvimento do SDW.

De notar que no decorrer da elaboração do plano estratégico para o desenvolvimento do SDW, foram também descortinadas possíveis oportunidades para o desenvolvimento de SI interdepartamentais que poderão ajudar a resolver alguns problemas relativos à integração e intercâmbio de dados diagnosticados no levantamento de requisitos.

Capítulo 3

Ciclo de Vida de um Projecto de um Sistema de *Data Warehousing*

3.1 Impacte de um SDW no Seio de uma Organização

3.1.1 Informação

A informação não é só por si reconhecida como conhecimento. A informação é a base para se obter conhecimento. No entanto, esta precisa de ser correcta e coerente. Não basta ter muita informação, é também necessário saber o que ela representa. Foi com base neste conceito que os gestores de empresas começaram a promover o desenvolvimento e a implementação de mecanismos para a análise dos dados fornecidos pelos seus sistemas transaccionais.

Normalmente, os sistemas transaccionais em funcionamento nas empresas têm como principal tarefa o suporte do dia-a-dia da organização, processando e armazenando os dados que são utilizados pelos diversos serviços de cada uma das secções de uma empresa, com o objectivo de automatizar os seus processos de tratamento de informação [Gupta '97].

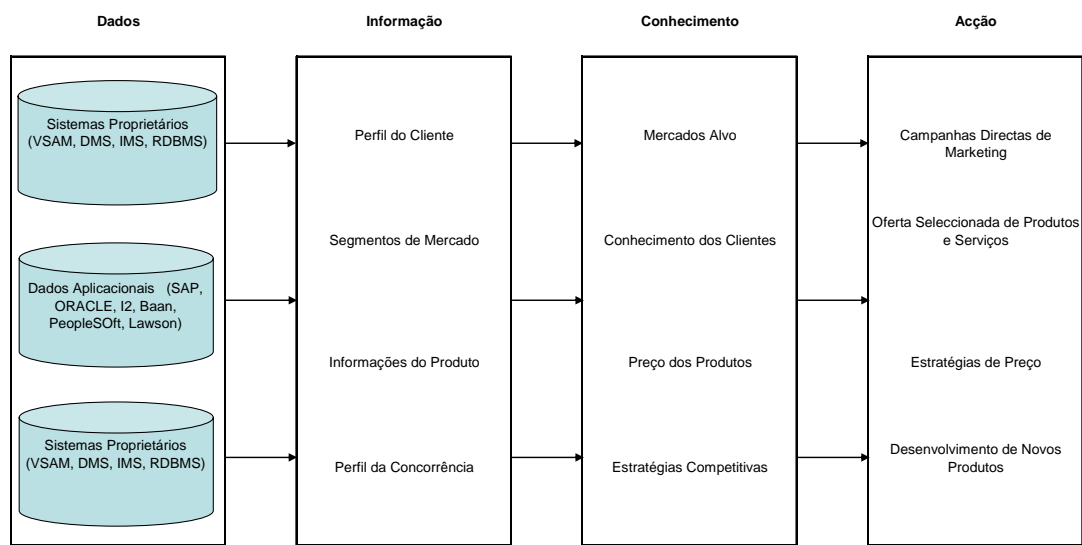


Figura 2 - Dados, Informação, Conhecimento, Acção

O crescimento do volume de informação gerado e tratado por este tipo de sistemas cresce de dia para dia, tornando-os mais lentos e exigindo cada vez mais espaço de armazenamento. Não é de surpreender assim, que a transferência dos dados mais antigos, e já sem qualquer influência nas actividades correntes da empresa, para outros sistemas de armazenamento secundário seja uma actividade comum, conseguindo-se desta forma atenuar em parte os problemas provocados por essa taxa de crescimento acentuada.

Porém, tal tarefa não é compatível com os requisitos exigidos pelos utilizadores mais credenciados dos sistemas que, normalmente, necessitam de muita dessa informação para poderem tomar as suas decisões. Este tipo de actividade de exploração de informação tão peculiar, especialmente orientada para assuntos de gestão bastante particulares e escolhendo "pacotes" de dados para

análise muito específicos, deu origem a uma nova vaga de sistemas: os sistemas de apoio à decisão.

3.1.2 Sistemas de Apoio à Decisão

Os sistemas de apoio à decisão têm evoluído conforme as necessidades dos seus utilizadores, forçando muitas vezes alguns departamentos ou unidades de serviços das organizações a criarem os seus próprios sistemas e armazenarem a informação conforme as suas necessidades específicas de gestão e de tomada de decisão. Em alguns casos, esta situação conduziu a que as organizações promovessem a criação de unidades de serviço vocacionadas especialmente para o projecto, implementação e manutenção de sistemas independentes, criados única e exclusivamente para responder às necessidades de informação de gestão e de tomada de decisão.

O contínuo desenvolvimento de sistemas de apoio à decisão com o intuito de resolver necessidades departamentais ou localizadas levou a que estes sistemas começassem a apresentar algumas deficiências. Da lista de deficiências conhecidas, as mais comuns relacionam-se normalmente com problemas acerca de dados incorrectos, mal definidos ou não partilhados, relatórios inconsistentes, e ausência de informação histórica acerca das actividades das próprias organizações.

A incompreensão dos dados existentes num sistema de apoio à decisão por parte de determinado departamento de uma organização não se reflecte tanto nos utilizadores actuais desse departamento, mas sim nos seus novos utilizadores ou noutros de diferentes departamentos. Isto pode-se justificar, pelo simples facto de que, na generalidade os sistemas são construídos tomando em consideração as necessidades dos utilizadores correntes, que, obviamente, lidam diariamente com esses dados e, acima de tudo, sabem como os interpretar. Esta é uma condição necessária e fundamental para se conseguir desenvolver processos efectivos de exploração e de análise de dados. A incompreensão dos dados armazenados num sistema de apoio à decisão é normalmente causada pela ausência ou falta de actualização da documentação existente que explique (e justifique) os valores tomados por determinadas estruturas de dados. Sem essa documentação, os

novos utilizadores têm o seu trabalho de análise um pouco mais complicado, já que ficam também com a tarefa, tantas vezes ingrata, de aprender com a experiência dos utilizadores mais veteranos.

A má definição dos dados provém também do problema dos sistemas de apoio à decisão serem desenvolvidos frequentemente tendo em conta apenas as questões de um departamento específico, desconsiderando em muitos casos a visão global da própria organização. A situação pode-se agravar mais um pouco quando estes sistemas são abertos a utilizadores de outros departamentos, já que a forma como estes vão abordar e interpretar os dados é muito diferente dos utilizadores desse departamento. Isto fará com que os utilizadores externos, desconhecedores da maior parte dos motivos que conduziram à situação actual dos dados desse departamento, tenham interpretações muito particulares, podendo conceber cenários de análise e avaliações de resultados completamente desfasados da realidade desse departamento. Como consequência dessas diferentes vistas sobre os dados, e do que estes na realidade representam, é normal que sucedam fortes discrepâncias entre relatórios de diferentes departamentos sobre o mesmo assunto, levando a que cada departamento considere o seu relatório como o correcto e o dos outros como errado, o que conduz, em muitos casos, a uma situação de mal-estar na organização com consequências óbvias para o futuro da própria organização.

Um dos resultados do mau desenvolvimento de sistemas operacionais e da má formação dos seus utilizadores é a introdução de dados de forma errada. Estes dados, uma vez transpostos para sistemas de apoio à decisão, podem levar os gestores a tomar decisões erradas. A tarefa de analisar frequentemente e com atenção os dados existentes nos sistemas transaccionais (e corrigi-los se necessário) é muito necessária, mas nem sempre é promovida ou mesmo executada. Fazer o diagnóstico e implementar a resolução deste tipo de problemas ajuda muito na futura implementação de qualquer sistema de apoio à decisão. No entanto, essa tarefa raramente é realizada, já que os técnicos possuidores de competência adequadas alegam, frequentemente, não terem disponibilidade para o fazerem. Dada a grande importância deste tipo de tarefa, é, no mínimo, um pouco estranho que isso possa acontecer.

Outra das consequências do desenvolvimento departamental de sistemas de apoio à decisão revela-se na falta de vontade na partilha da informação produzida. Uma vez que estes sistemas

foram desenvolvidos dentro do departamento pelo seu pessoal técnico - que têm absoluto controle sobre o seu desenvolvimento e manutenção - recorrendo a verbas próprias, é muito normal que a partilha de dados ou o controle do sistema não seja realmente muito encorajada. Todavia, mesmo quando existe essa vontade de partilhar dados e sistemas, essa tarefa é de muito difícil realização. A maioria das organizações não tem ainda os seus sistemas integrados ou com a possibilidade de suportar processos de cruzamento de dados, apesar de estarmos a assistir actualmente a um esforço considerável das organizações nesse sentido [Adelman et al. '00].

3.1.3 Sistemas de *Data Warehousing*

O desenvolvimento de um SDW tem como principal objectivo colmatar as deficiências dos sistemas de apoio à decisão e a inadequação dos sistemas transaccionais para responder a questões relacionadas com as actividades de gestão e de tomada de decisão das empresas. As principais razões que levam as empresas a implementar um SDW são, nomeadamente, a necessidade de [Greenfield '02b]:

- Executar tarefas associadas com a resposta a questões e elaboração de relatórios em servidores não utilizados pelos sistemas transaccionais.
- Fornecer um ambiente no qual seja necessário pouco conhecimento sobre as bases de dados existentes para a elaboração e manutenção de questões e de relatórios pelos utilizadores.
- Utilizar modelos de dados e tecnologias orientadas à resolução de questões e à elaboração de relatórios.
- Fornecer um repositório de dados tratados que possa ser usado para responder a questões ou para elaborar relatórios.
- Facilitar a resposta a questões e elaboração de relatórios sobre dados provenientes de diversos sistemas e até de dados em armazenamento secundário.
- Prevenir o acesso aos sistemas transaccionais de utilizadores que apenas necessitem de respostas a questões e de elaborar relatórios sobre assuntos relacionados com actividades operacionais.

No entanto, a implementação destes sistemas não é uma tarefa fácil nem deve ser encarada como a implementação de mais um projecto informático. A construção de um SDW é mais abrangente que o mero desenvolvimento de uma aplicação informática, uma vez que tem como fontes de informação elementos muito diferentes, tais como sistemas transaccionais, folhas de cálculo ou ficheiros de texto. Um SDW tem como objectivo essencial a integração da informação proveniente dessas fontes num só repositório de dados, que possa ser depois acedido e questionado pelos utilizadores que estejam envolvidos com questões relacionados com a gestão e tomada de decisão da organização. A existência de um sistema destes é, hoje em dia, vital para o desenvolvimento de uma organização. Assim, para que alcance uma situação de sucesso, é crucial que seja feito um planeamento cuidadoso e um controlo muito apurado no desenvolvimento deste tipo de projectos.

3.2 Sucesso e Insucesso de um SDW

A probabilidade de insucesso de um SDW continua a ser um dos factores que classifica este tipo de projectos como sendo de alto risco, não só pela dificuldade de implementação, mas também pelos custos envolvidos, sejam estes relacionados com recursos tecnológicos, humanos ou temporais [Srivastava et al. '99]. Uma das formas de minimizar essa probabilidade é através da identificação dos diversos factores de risco. Estes terão que ser bem geridos para prevenir o insucesso e atingir os objectivos inicialmente previstos para o SDW.

Entre os diversos factores de risco existentes num projecto de implementação de um SDW existem alguns que são especialmente delicados e que influenciam de forma vinculada o futuro do sistema. Por exemplo, a gestão da expectativas, o envolvimento dos utilizadores no projecto, a existência de um bom patrocinador, uma equipa técnica competente, um planeamento temporal realista, um orçamento equilibrado, são alguns dos factores (tecnológicos) inerentes ao processo de implementação do SDW que devem ser tomados sempre em consideração, de forma muito séria [Adelman et al. '00].

Uma boa ou má gestão das expectativas do projecto pode determinar o sucesso ou o insucesso da implementação de um SDW. É aconselhável que os utilizadores tenham uma ideia bastante clara do que é na realidade um SDW e como é que este os poderá ajudar nas suas actividades quotidianas. Concretamente, os utilizadores deverão saber muito bem o tipo de contributos que um SDW pode dar e conhecer o que é necessário fazer para que tais contributos possam ser úteis. Por exemplo, uma simples questão efectuada sobre tabelas de grandes dimensões poderá não ter uma resposta muito rápida, podendo demorar alguns minutos ou até algumas horas. Os utilizadores deverão estar sensibilizados para esse tipo de situação.

Os utilizadores deverão saber quais as funcionalidades do sistema, quais os dados que podem ser acedidos, quais as questões e relatórios predefinidos no sistema e qual o nível de detalhe existente no SDW. Complementarmente deverão também estar informados relativamente à disponibilidade do sistema. Assim, deverá ser comunicado a todos eles os dias, e em que horário, é que o sistema estará disponível.

Uma vez que um SDW serve de suporte à análise dos dados – tendências e variações –, os utilizadores deverão saber qual a informação que está armazenada no sistema, já que este tipo de análises exige, normalmente, conhecimento acerca de dados históricos relacionados com períodos de actividade passada das empresas. Assim é de todo conveniente informá-los acerca do estado dessa informação e sobre o tempo que permanecerá no sistema até que seja transferida para um sistema de armazenamento secundário.

O desenvolvimento de processos de formação adequados para todos os utilizadores do sistema é essencial para a boa aceitação do SDW. A implementação do sistema está muitas vezes condicionada pela participação e pelo tipo de conhecimento que os utilizadores têm acerca do sistema. Os futuros utilizadores do SDW deverão também ter conhecimento dos prazos previamente estabelecidos para a conclusão do projecto. Como um SDW é, normalmente, desenvolvido passo a passo, de forma iterativa, os utilizadores não terão acesso a todas as funcionalidades do sistema ao mesmo tempo. Deverá, pois, ser estabelecido um prazo para a conclusão de cada iteração, sendo necessário para cada uma delas enunciar quais os utilizadores que passarão a ter acesso ao sistema após a conclusão de cada uma das iterações.

Assim que os novos utilizadores comecem a utilizar o novo sistema muitas dúvidas serão levantadas relativamente às suas funcionalidades e modelos de operação. Sendo assim, deverá ser criada uma equipa especialmente orientada para suporte às actividades dos utilizadores. Essa equipa deverá ter as competências necessárias para responder às diversas dificuldades comunicadas pelos utilizadores no âmbito de cada uma das áreas funcionais dos vários departamentos abrangidos pelo projecto.

Todos estes factores estão relacionados com as expectativas que são criadas pelos utilizadores quanto ao SDW e que deverão ser abordados num documento específico durante a fase de elaboração do projecto que será submetido à aceitação dos utilizadores. Além da gestão dessas expectativas existem outros factores, bastante críticos, que influenciam o sucesso da implementação de um projecto de um SDW. De referir:

- o envolvimento dos utilizadores no desenvolvimento do projecto;
- a escolha de um patrocinador do projecto;
- a escolha dos membros da equipa de planeamento e desenvolvimento do projecto;
- a qualidade dos dados dos sistemas fonte;
- o calendário das tarefas a executar;
- a elaboração de um orçamento adequado;
- a escolha das ferramentas a utilizar no desenvolvimento do projecto;
- a arquitectura e desenho do SDW e a integração das diversas plataformas necessárias;
- o abandono de elementos chave para o projecto;
- a implementação de um SDW distribuído;
- a formação dos utilizadores.

O envolvimento dos utilizadores no projecto de um SDW é assim um elemento chave para o sucesso. Os utilizadores aderirão mais facilmente a um sistema do qual fizeram parte do que a um sistema para qual não contribuíram.

A escolha do patrocinador do projecto é também um factor crítico, devendo este ser bastante conhecedor da área do negócio do projecto e não da sua área tecnológica. O seu papel principal será o de garantir todas as condições necessárias, sejam elas físicas ou financeiras, para a conclusão com sucesso do projecto, agindo como elemento desbloqueador de possíveis problemas entre a equipa de desenvolvimento e os outros departamentos da organização envolvidos. Quanto à equipa de planeamento e desenvolvimento do projecto esta deverá ser constituída por elementos que sejam especialistas nas diversas áreas necessárias ao projecto. No entanto, poderá haver para uma determinada tarefa vários especialistas e vice-versa.

A qualidade dos dados fonte é um factor imprevisível. Uma má definição e implementação dos sistemas transaccionais poderá originar grandes quantidades de dados mal definidos ou incorrectos. Tal facto irá conferir, com certeza, à tarefa de extracção e limpeza de dados um factor de risco muito elevado, que irá consumir muito tempo ao projecto, podendo inclusive atrasá-lo ao ponto de se tornar uma tarefa quase impossível [Lee et al. '99, Maletic et al. '00].

A elaboração de um calendário de tarefas realista é também uma ajuda para a implementação com sucesso do projecto. A falta de experiência de alguns intervenientes do projecto leva a criação de um calendário irrealista e, consequentemente, a um orçamento inadequado, podendo originar, caso haja atrasos, a perda de patrocínio e falta de apoio, atrasando ainda mais o projecto e condenando-o ao insucesso.

A escolha das ferramentas necessárias à implementação do SDW é outro factor de risco, dado o elevado custo de aquisição e manutenção dessas ferramentas. Estas devem ser previamente avaliadas e ajustadas aos requisitos do projecto, sem que para tal se perca muito tempo, uma vez que o tempo é, por si só, um bem escasso. Por outro lado, a integração de plataformas heterogéneas no SDW também pode levantar algumas questões críticas, uma vez que eleva o nível de dificuldade de implementação do projecto, em particular na fase de extracção e transformação dos dados [Hernández et al. '95]. A implementação de um SDW distribuído eleva também, consideravelmente, o risco do projecto, dada a complexidade de implementação de sistemas distribuídos ser normalmente bastante elevada.

A perda de um elemento da equipa de desenvolvimento do projecto também pode ser bastante problemática, especialmente se esse elemento for responsável pela realização de várias tarefas e se a equipa não integrar outro elemento com o mesmo tipo de conhecimento e perícia. Por outro lado, a ausência de processo de formação ou a eventual formação inadequada dos elementos da equipa de desenvolvimento ou dos utilizadores finais do sistema são questões muito delicadas que também podem elevar significativamente o risco de insucesso do projecto.

Apesar de já terem sido referidos muitos dos factores de riscos e aspectos que podem contribuir para o insucesso de um SDW, ainda existem alguns outros que devem ser tomados em consideração por qualquer elemento responsável pela implementação de um SDW, em particular quando:

- o projecto ainda na fase de desenvolvimento esgotou o seu orçamento;
- os utilizadores encontram-se insatisfeitos com a qualidade dos dados e com as ferramentas de acesso ao SDW após a implementação do projecto;
- o fraco desempenho do sistema e a fraca adesão dos utilizadores ao novo sistema prejudica a sua implementação na organização;
- o processo de extracção, transformação e carregamento dos dados no SDW consome mais tempo que a janela da oportunidade.

Decorrido algum tempo sobre a implementação do SDW é possível retirar algumas conclusões e quantificar o impacte do projecto na organização. A avaliação mais comum costuma ser na análise do retorno do investimento. Se usando o SDW houver uma redução de custos, através de melhor gestão de stocks, se houver um acréscimo de produtividade, através de uma melhoria na gestão de recursos humanos, ou se houver um acréscimo de receitas, através da identificação de novos nichos de mercado ou de novos produtos, o SDW contribuiu assim para o retorno do investimento efectuado no seu desenvolvimento. É assim possível medir o sucesso da sua implementação.

Um projecto de um SDW pode também ser considerado um sucesso se este respeitar os pressupostos descritos no seu planeamento, nomeadamente em relação à calendarização prevista, ao orçamento previsto, às metas previstas, através da resolução de alguns problemas do negócio

ou através da identificação de novas oportunidades de negócio. Outras medidas para quantificar o sucesso são o cumprimento dos objectivos iniciais, a análise da utilização efectiva do *data warehouse*, a satisfação dos utilizadores, e se essa satisfação se traduz num acréscimo de pedidos de utilização do SDW.

3.3 O Projecto de um SDW

A construção de um SDW como um sistema monolítico é normalmente um passo para o insucesso do sistema, daí que a melhor aproximação seja uma implementação gradual, de forma iterativa [Gonzales '03, Kimball et al. '98]. No entanto, o primeiro passo, a primeira iteração, é a mais complexa por ser a base para todas as outras iterações. O desenvolvimento de um *data warehouse* é um processo evolutivo, que pode ser encarado como uma espiral, em que as iterações seguintes começam sempre tendo como base a conclusão das iterações anteriores. A adaptação do modelo em espiral, proposto por Boehm, ao desenvolvimento de um SDW, transforma cada iteração de um projecto de *data warehousing* num processo organizado em cinco Fases [Craig et al. '99]: análise, planeamento, desenho, desenvolvimento, instalação e formação.



Figura 3 - Modelo de Desenvolvimento de *Software* (adaptado de [Boehm '88])

Cada uma destas fases é constituída por um conjunto de tarefas que devem ser desempenhadas ou analisadas conforme seja a primeira iteração ou uma das seguintes.

3.3.1 Análise

A fase de Análise é compreendida por um conjunto de tarefas entre as quais se encontra a análise da necessidade de implementação de um *data warehouse* numa empresa, isto é, basicamente, a elaboração do seu projecto tendo como base uma análise ponderada dos requisitos orientados ao negócio. Esta análise deverá entrar com vários factores e respectivas ponderações de forma a obter uma lista de requisitos ordenada pela importância e impacte que terão no futuro da empresa. De notar que este documento deverá ser revisto e as ponderações recalculadas no início de cada iteração, uma vez que as prioridades de hoje poderão não ser as prioridades de amanhã.

Nesta fase deve ser escolhido o Gestor do Projecto que será responsável pela liderança de todo o projecto, por cada uma das suas fases e pela interacção com todos os intervenientes externos ao projecto. Como garantia de apoio ao projecto deve também ser escolhido um ou mais patrocinadores da empresa que garantam que este terá todas as condições necessárias para o seu desenrolar. De referir, em particular, os recursos humanos e financeiros que permitam a conclusão com sucesso do projecto. Estes, têm também a tarefa de facilitar a comunicação entre o Gestor do Projecto e as diversas secções da Empresa.

É também nesta fase que deve ser elaborado um documento de auto-avaliação, de forma a determinar se a empresa está ou não preparada para embarcar num projecto desta natureza, a definir os pontos fortes e fracos da empresa que poderão influenciar a conclusão com sucesso do projecto de *data warehousing*, e a avaliar o impacte do SDW nos actuais sistemas transaccionais e políticas organizacionais da empresa. Nesta fase de análise de requisitos devem estar envolvidas pessoas com diferentes perfis, tais como:

- **Patrocinador**, que integrará, em princípio, gestores com poderes executivos dentro da organização que definem quais os objectivos e prazos de implementação e que garantem que o projecto tem os recursos, quer financeiros quer humanos, necessários à sua implementação.
- **Gestor do projecto**, que é o responsável máximo pela elaboração, gestão e implementação do projecto do SDW baseado nas necessidades do negócio.
- **Analista do negócio**, que integra os elementos do projecto que são normalmente consultores externos especializados nas diversas áreas de negócio em que se posiciona a organização; têm como responsabilidade a análise e projecção da evolução do negócio com vista a uma melhor adaptação ao futuro da organização em questão.
- **Especialista dos sistemas transaccionais**, que é um perfil que reúne normalmente elementos pertencentes aos departamentos de informática das empresa e que tem conhecimento sobre as tarefas desempenhadas pelos sistemas transaccionais, assim como dos dados e dos esquemas de bases de dados neles armazenados.
- **Administrador de bases de dados**, que é constituído pelos responsáveis pelo desenho das bases de dados, implementação, controlo do desempenho e procedimentos de segurança dos sistemas de desenvolvimento e de produção.

3.3.2 Planeamento

Após a conclusão da primeira fase, e tendo como base o projecto delineado, cabe ao gestor do projecto elaborar um documento com os recursos humanos necessários para cada tarefa, especificando o tempo necessário de cada recurso. Nesse documento deve também ser especificado quais os recursos tecnológicos necessários à implementação do SDW. A documentação de um projecto de um SDW é uma questão fundamental. A documentação gerada deve ser detalhada e suficientemente abrangente, cobrindo todos os aspectos relacionados directa ou indirectamente relacionados com o projecto. Sobre a documentação gerada deverão ser aplicados critérios de gestão bastante efectivos, de forma a assegurar a sua manutenção e difusão.

Anteriormente, já foi referida a necessidade de se elaborar documentos relativos aos requisitos do negócio, às necessidades em termos de tempo e recursos relativos à implementação do sistema. Agora, é necessário também, no decorrer de todo o projecto, documentar todas as ocorrências – datas, tomadas de decisão ou resultados – para um melhor entendimento e análise futura do projecto do SDW. Durante esta fase, devemos também elaborar, ou se necessário rever, um documento com uma visão mais alargada daquilo que se pretende alcançar com a implementação do SDW. Esta visão de longo prazo dar-nos-á algumas indicações importantes, como por exemplo, informação necessária ao suporte das decisões relativas à implementação tecnológica do SDW. Contudo, a visão a curto prazo é também muito relevante no âmbito do projecto e, como tal, deve ser contemplada na documentação gerada, tendo em consideração o documento que foi previamente elaborado contendo a relação dos recursos necessários à implementação desta iteração e respectivo calendário de tarefas a desenvolver no âmbito do projecto.

Outra questão a ter em conta, principalmente durante a fase de planeamento do projecto, apesar de não se dever descurá-la nas restantes fases, é a comunicação entre a equipa que está a desenvolver o SDW, os utilizadores e os patrocinadores do projecto. É vital para o sucesso do projecto que todos se sintam envolvidos no seu desenvolvimento e que participem activamente, sempre de forma construtiva. Se todos sentirem que o SDW vai melhorar a forma como desempenham o seu trabalho, contribuindo consequentemente para um melhor funcionamento da empresa, então todos estarão a “remar” no mesmo sentido e a probabilidade de sucesso do projecto aumenta significativamente. O desenvolvimento de um SDW à margem dos utilizadores tende a ser um sistema incompreendido, pouco utilizado e condenado ao fracasso. Este é um aspecto crucial no desenvolvimento de um SDW. Todos os factores de risco identificados devem ser devidamente documentados e tratados com a maior celeridade possível, de forma a que o projecto não se atrase nem se desvie do seu objectivo. Caso contrário estaremos a aumentar a probabilidade de insucesso do Projecto.

Durante a fase de planeamento a equipa do projecto poderá integrar, para além dos perfis de gestor do projecto, especialista dos sistemas transaccionais e administrador de bases de dados, apresentados já durante a apresentação da fase de análise, elementos com os seguintes perfis:

- **Analista técnicos**, que integra os elementos que têm como principal tarefa a análise das infra-estruturas tecnológicas da organização, sendo-lhes atribuída, normalmente, a tarefa de elaborar o desenho da futura arquitectura tecnológica que irá suportar o SDW.
- **Analista de dados**, perfil cujos elementos têm como principal tarefa a interpretação dos dados presentes nas diversas fontes para que no futuro possam elaborar os planos de transformação, limpeza e integração de dados no SDW.

3.3.3 Desenho

A base de trabalho para esta fase é o documento com os requisitos do sistema, definidos em parte pelos utilizadores futuros do sistema. Nessa análise foi determinada qual a próxima iteração, daí que, caso estejamos no início do projecto, deveremos desenvolver um protótipo como forma, não só de garantir o apoio dos utilizadores como dos patrocinadores, como também para ficarmos com uma maior certeza do rumo correcto do desenvolvimento.

Na fase de Desenho deve ser elaborado o desenho lógico da base de dados, sempre em conformidade com anteriores iterações. Deve também ser elaborado ou revisto o desenho das ferramentas de *Extracção, Transformação e Integração* (ETI) de dados. A arquitectura tecnológica deve também ser definida, ou revista, conforme estejamos na primeira iteração ou nas seguintes. Para finalizar, deve também ser elaborado o desenho das aplicações que vão ser usadas pelos utilizadores finais. Nesta fase a equipa será constituída por um ou mais elementos com os seguintes perfis: gestor do projecto, especialista dos sistemas transaccionais, administrador de bases de dados, analista técnicos, analista de dados. Todos estes perfis já foram descritos anteriormente durante a apresentação das fases de análise e planeamento.

3.3.4 Desenvolvimento

A fase de desenvolvimento tem como objectivo principal a implementação de todas as tarefas que foram especificadas durante a fase de desenho. Caso seja possível, e no sentido de minimizar o

tempo necessário ao desenvolvimento do projecto, deve ser ponderada a hipótese do desenvolvimento ser realizado em paralelo segundo quatro áreas diferentes: base de dados, ferramentas ETI, arquitectura tecnológica e aplicações. Nesta fase a equipa deverá ser constituída por um ou mais elementos com perfis de gestor do projecto, especialista dos sistemas transaccionais, administrador de bases de dados, analista técnicos, analista de dados programadores e especialistas em suporte técnico. Todos estes perfis já foram caracterizados em fases anteriores à excepção dos dois últimos perfis referidos. Estes têm as seguintes caracterizações:

- **Programador**, que integra os vários elementos do projecto cuja tarefa principal é a exploração e implementação das ferramentas ETI e das aplicações que vão interagir com o SDW, tal como foram definidas durante a fase de desenho do projecto.
- **Especialistas em suporte técnico**, que é o perfil no qual estão definidos os responsáveis por resolver os pequenos problemas que vão surgindo com o desenvolvimento do SDW, principalmente aqueles que afectam os sistemas transaccionais e o SDW, caso já não estejamos na primeira iteração.

3.3.5 Instalação e Formação

Como conclusão de uma iteração, deve ser feita a instalação de todo o *hardware* e *software* necessário à entrega do sistema desenvolvido e dada a formação necessária a todos os utilizadores que irão interactivar com o sistema por forma a se garantir a maximização da utilização e das potencialidades do novo sistema. Só assim poderemos garantir a implementação de um SDW com sucesso. Nesta fase do projecto a equipa deverá ser constituída, para além do gestor de projecto, por um ou mais elementos com o perfil de formador. Os elementos com este perfil serão os responsáveis pela correcta formação dos utilizadores neste novo ambiente, nomeadamente através do uso das aplicações que interagem com o SDW e os ajudarão na interpretação dos dados. No início do projecto há também a necessidade de alguns elementos da equipa receberem formação em diversas áreas – por exemplo modelação dimensional, SQL, *reporting*, entre outras - para que estes estejam preparados para o projecto.

3.4 O Dia Seguinte ao Arranque do SDW

A entrada em funcionamento do SDW não determina, só por si, o sucesso do projecto, existe um número elevado de questões e tarefas a desempenhar que vão determinar o seu sucesso a longo prazo. Foi referido anteriormente a constituição de cada uma das iterações de um projecto de um SDW. Nessa constituição eram focadas cinco fases distintas, nomeadamente análise, planeamento, desenho, desenvolvimento e, finalmente, a fase de instalação e formação. Uma vez que um SDW é um sistema evolutivo, de desenvolvimento iterativo, estas fases repetem-se em cada nova iteração que sucede. No entanto, existe um conjunto de tarefas que só aparecem após a conclusão da primeira iteração e que vão coexistir com o desenvolvimento do sistema. Poderemos chamar a esse conjunto de tarefas a fase de pós-produção [Adelman et al. '98, Kimball et al. '98, Westerman '01]. Esta fase pertence ao ciclo de vida de um projecto de um SDW, surge na conclusão da primeira iteração e funcionará em paralelo com as iterações seguintes.

As questões e tarefas a abordar na fase de Pós-Produção estão relacionadas com um conjunto variado de temas, são eles: refrescamento dos dados, desempenho do sistema, evolução das fontes de dados, utilizadores. As tarefas e questões relacionadas com o refrescamento dos dados têm como foco principal a diferença entre o método utilizado para o primeiro carregamento da base de dados e os carregamentos seguintes. A tarefa de determinar quais os dados a serem carregados em cada refrescamento deverá ser executada com muito cuidado e atenção.

Se o projecto do SDW foi bem elaborado, a arquitectura do SDW será, naturalmente, escalável, quer em termos de desempenho quer em termos de armazenamento. No entanto, é necessário na fase de pós-produção avaliar frequentemente o desempenho do SDW e se este está ou não dentro dos critérios estabelecidos inicialmente com os seus utilizadores e gestores. Se tal não acontecer, deve-se considerar uma actualização do sistema de forma a melhorar o seu desempenho ou a sua capacidade de armazenamento.

Outro problema a ter em atenção é o facto das fontes de dados estarem em permanente evolução. Partindo do princípio que a maior parte dos dados provêm de sistemas transaccionais, que o negócio está em permanente evolução e que as necessidades de representação e armazenamento de dados também, é necessário fazer-se uma permanente monitorização dos dados extraídos desses sistemas e uma análise cuidada sobre o seu significado.

Se o SDW tiver aceitação por parte dos utilizadores, estes farão dia após dia cada vez mais solicitações o que afectará o desempenho do sistema. Também surgirão pedidos de novas funcionalidades que terão de ser satisfeitas por parte da equipa de pós-produção. O aparecimento de novos utilizadores será então uma consequência natural, trazendo eles também novos requisitos e pedidos de funcionalidades.

De referir, por fim, que a aceitação do SDW por parte dos utilizadores é claramente um factor de sucesso. Todavia, este facto pode também condenar um projecto ao insucesso caso não haja a resposta adequada do SDW ou da equipa de pós-produção na análise e implementação desses novos pedidos.

Capítulo 4

Kimball, Inmon e Outros

4.1 A Abordagem de William H. Inmon

William H. Inmon é considerado pela comunidade informática como o pai do conceito de *Data Warehousing*. Inmon tem vindo a apresentar a sua visão sobre a metodologia a adoptar no desenvolvimento de SDW [Inmon '94, '97, '02]. Com o evoluir do tempo, a sua metodologia tem vindo a ajustar-se às necessidades do mercado colmatando algumas falhas verificadas na sua primeira abordagem. Nas próximas secções descreve-se a visão de Inmon no desenvolvimento de SDW, focando com maior atenção a vertente do Projecto, da Gestão e da Implementação do SDW.

4.1.1 Projecto

Inmon defende que a construção de um SDW como uma só tarefa acarreta demasiados riscos e, como tal, está destinada ao insucesso. Daí que, o desenvolvimento de um SDW, deva ser

abordado através da criação de um plano onde, iterativamente, passo a passo, são sucessivamente adicionados novos requisitos.

Adicionalmente, Inmon define basicamente um projecto de *data warehousing* como um plano para a migração de dados entre os sistemas operacionais e um repositório especializado para suporte à decisão designado por *data warehouse*. A migração do dados para o *data warehouse* não é directa, envolve muitas operações intermédias de tratamento, tais como limpeza, conversão, filtragem ou substituição de dados. O plano de migração referido assenta num modelo de dados corporativo, que deve representar as necessidades da organização, identificar quais os principais temas existentes na organização, a sua definição e respectivos relacionamentos. Este modelo de dados deve também identificar quais as chaves e os atributos que representam cada um desses temas e identificar eventuais grupos de caracterização repetidos.

Uma vez terminado o modelo de dados corporativo, deve construir-se um modelo de dados específico a cada tema, ou seja, um modelo um pouco mais refinado, mais próximo dos elementos base do sistema. Após construção dos modelos de dados que definem os requisitos da organização, quer a alto nível quer a nível intermédio, é necessário identificar quais as fontes de informação que vão fornecer os dados para alimentar o modelo de dados construído. Para isso, deve-se escolher as “melhores” fontes de informação, avaliando-as segundo um conjunto de critérios específicos – disponibilidade, qualidade, acessibilidade, operacionalidade, proximidade ou desempenho dos sistemas. O processo de avaliação deve analisar não só os sistemas de suporte dessas fontes de informação como também, e prioritariamente, os dados que elas contêm, já que são eles que darão sentido à própria construção do *data warehouse*. Assim, deverá ser realizada com rigor uma análise detalhada sobre os dados que estão disponíveis, tendo em consideração aspectos como a sua completude, correcção, precisão, consistência, redundância, acesso e é claro a sua própria caracterização – os metadados.

O passo seguinte consiste no desenho do próprio *data warehouse* e da definição dos processos de transformações dos dados que foram extraídos dos sistemas operacionais e que destinam ao novo sistema. Uma vez concluído este passo, deve-se desenhar e desenvolver o interface entre os

sistemas operacionais e o *data warehouse*, isto é, os programas e as estruturas de dados que suportarão os processos de extracção, transformação e integração dos dados no *data warehouse*.

Há quem defenda que a construção de um *data warehouse* só seja feita quando os sistemas operacionais estiverem limpos. Esta ideia não é defendida por Inmon, uma vez que tal atitude apenas levará a que nunca se consiga desenvolver um SDW.

Uma vez alimentado – povoado – o *data warehouse* está então concluída a primeira fase do processo, a primeira iteração. Agora os utilizadores finais já podem explorar os dados armazenados no *data warehouse* durante os seus próprios processos de análise e pesquisa. Com esta exploração ter-se-ão as primeiras reacções à utilização do sistema que funcionarão como uma espécie de “feedback” na comunicação de eventuais requisitos de informação ou alterações funcionais à equipa de desenvolvimento do SDW, que, por sua vez, reajustará, dentro das suas competências, o SDW de forma a que este possa responder a essas novas solicitações. A comunicação destas reacções é vital na vida de um SDW, já que prova a sua utilidade para a empresa e garante que este esteja sempre ajustado às reais necessidades da organização.

4.1.2 Gestão

Uma vez que a metodologia de desenvolvimento de SDW de Inmon é centrada nos dados, ou seja, é tipicamente uma abordagem *Top-Down* na qual é definida em primeiro lugar o modelo de dados corporativo do sistema de *data warehousing*, depois as ferramentas ETI e só depois as aplicações de consulta e análise. A componente de gestão do projecto é pouco abordada neste processo. No entanto, Inmon refere que uma das técnicas mais eficientes para garantir a qualidade nestes projectos é através da elaboração de uma *checklist* a rever no final do desenho físico do tema escolhido a adicionar ao *data warehouse*. Os elementos que deverão participar nesta revisão serão todos aqueles que estão envolvidos no projecto, em particular, o administrador da base de dados, os programadores, os analistas dos sistemas de apoio à decisão, os utilizadores finais, o pessoal de suporte técnico e, naturalmente, a gerência da empresa.

Para que este processo de revisão seja mais efectivo são promovidas reuniões que têm como principal objectivo resolver, debater, solucionar qualquer problema que surja no decorrer do desenvolvimento do SDW. Como sugestão, Inmon defende que a liderança dessas reuniões deverá ser feita por um elemento externo ao projecto, com o objectivo de se atingir a isenção desejada e poder-se obter uma nova perspectiva sobre o projecto. Em [Inmon '02] é proposto uma *checklist* modelo a usar em cada uma dessas reuniões. Nessa *checklist* são considerados assuntos que poderão de alguma forma comprometer o sucesso do projecto, como questões relacionadas com o desenho da base de dados, desenvolvimento do *data warehouse* e gestão do projecto.

4.1.3 Implementação

Inmon propõe uma metodologia de desenvolvimento orientada aos dados constituída por quatro fases distintas, nas quais se descrevem as actividades específicas assim como os resultados esperados. A saber:

- **METH1**, que descreve o processo de desenvolvimento dos sistemas operacionais.
- **METH2**, que aborda a implementação do SDW.
- **METH3**, que descreve o desenho e especificação dos processos de acesso ao conteúdo do *data warehouse*.
- **METH4**, que está associada com as tarefas de desenvolvimento relacionadas com os processos definidos na fase anterior

Como podemos observar na Figura 4, a implementação de um SDW tem como base o modelo corporativo de dados (DSS1). Caso este ainda não esteja construído terá de o ser obrigatoriamente, antes de se prosseguir para a etapa seguinte. A implementação tem também como base a análise dos requisitos tecnológicos previamente elaborados.

Com base no modelo corporativo de dados é então possível estimar qual o volume de informação projectado para o *data warehouse* (DSS2) e, como tal, tomar as devidas precauções. Isto significa que, caso o volume de dados seja muito grande deveremos entrar em consideração com diferentes níveis de granulosidade. Esta tarefa é executada uma vez por cada uma das iterações do sistema e

costuma consumir entre dois dias a duas semanas, consoante a definição do modelo de dados e do âmbito do projecto.

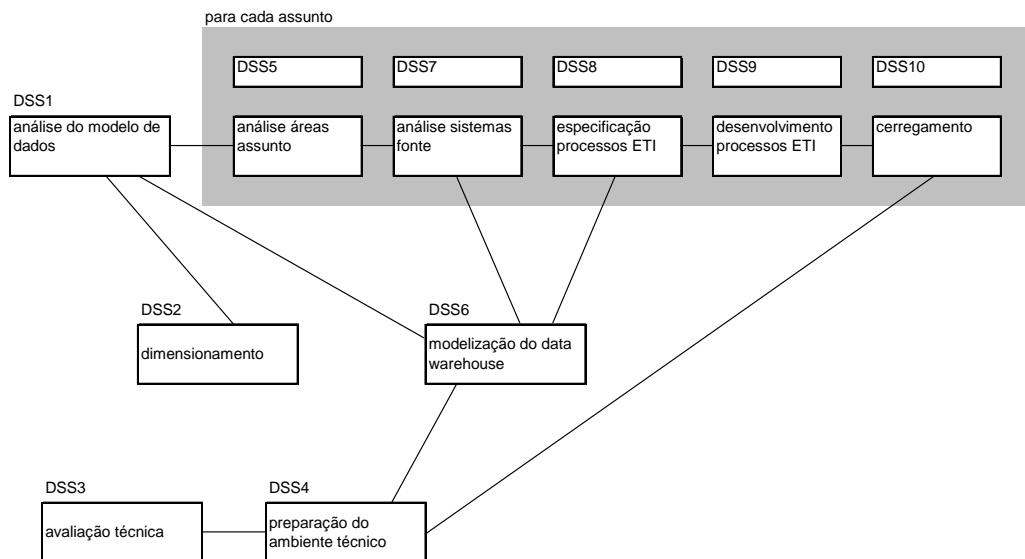


Figura 4 - METH2 [Inmon '02]

Ao mesmo tempo, e com base na análise dos requisitos tecnológicos, deve ser preparada a arquitectura tecnológica de suporte de acordo com os requisitos estipulados para o *data warehouse* (DSS4), que inclui a definição, a construção ou a adaptação da rede informática, a definição e aquisição do sistema de gestão para o *data warehouse*, a construção do próprio *data warehouse*, a definição e construção do interface com o *data warehouse*. Esta tarefa é executada uma vez e dura entre uma semana a um mês.

De seguida, deve ser escolhido um tema para ser implementado (DSS5). Este deve ser abrangente para que a sua implementação tenha algum tipo de impacte na organização. Todavia, não deve ser demasiado abrangente que torne o seu desenvolvimento uma tarefa impossível. Esta tarefa é executada uma vez por cada tema e consome um dia. Na tarefa DSS7 será necessário analisar os sistemas fonte e identificar qual a origem dos dados que irão ser transferidos para o *data warehouse* e, uma vez que poderão existir diferentes sistemas operacionais com a mesma informação, será necessário escolher qual a fonte mais credível, disponível, com informação mais

precisa e cujos dados sejam de fácil extracção. Esta tarefa é executada uma vez por cada tema e consome cerca de uma semana de trabalho.

O desenho da base de dados para o *data warehouse* apenas pode ser feito depois de definido o modelo de dados corporativo, que é independente de todas as considerações físicas, definida a arquitectura tecnológica, a grande condicionante para a concretização desta tarefa, e identificados os sistemas fontes que albergam os dados pretendidos. Esta tarefa (DSS6) é executada apenas uma vez e poderá demorar entre uma a três semanas para ser concluída.

O passo seguinte (DSS8) é a definição das ferramentas que irão ser responsáveis pelas operações de extracção, transformação e integração dos dados provenientes das diversas fontes de informação incluídas no projecto. A periodicidade desta tarefa é, mais ou menos, uma vez por cada ferramenta que seja necessária desenvolver. Depois de realizada a especificação das diversas ferramentas necessárias passa-se para o seu estudo e consequente desenvolvimento (DSS9). Tal como na tarefa anterior, esta deverá ser executada semanalmente por cada ferramenta escolhida.

Para finalizar temos a tarefa DSS10. Esta desenvolve-se em torno do carregamento (integração) dos dados no *data warehouse* utilizando as ferramentas previamente desenvolvidas durante a fase anterior. Quando executada com sucesso, esta tarefa contribuirá para a disponibilização de um *data warehouse* compreensível e acessível a toda a comunidade de utilizadores de uma empresa envolvidos em actividades de gestão e de tomada de decisão.

Uma vez alimentado o *data warehouse*, é necessário depois especificar e desenvolver as ferramentas que irão interagir com o SDW. Na fase METH3 são especificadas as ferramentas que irão permitir tirar partido dos conteúdos armazenados no *data warehouse*. A fase METH4 tratará de desenvolver essas mesmas ferramentas.

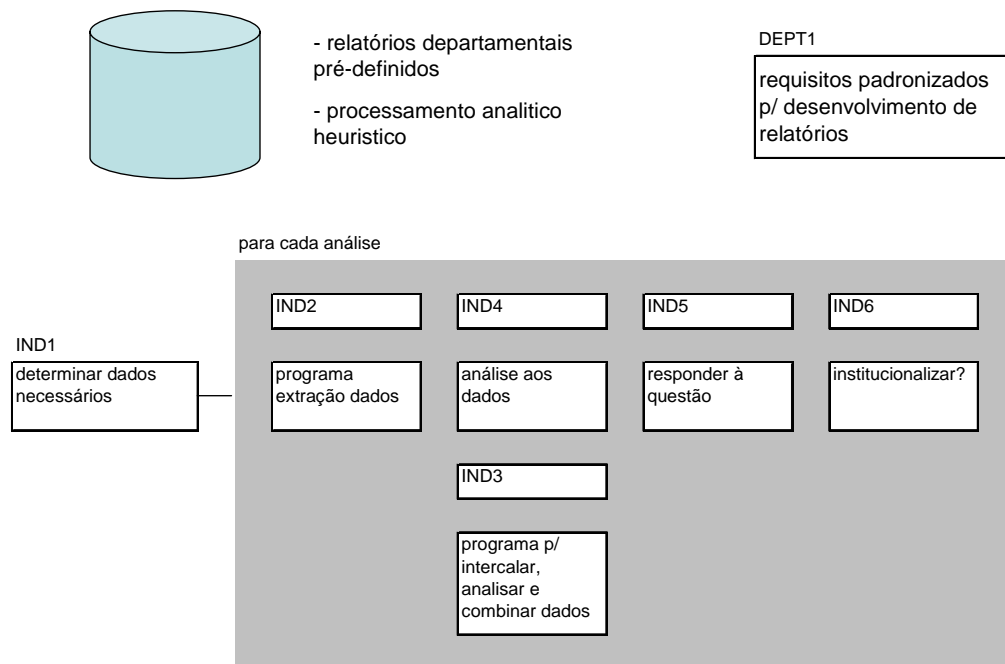


Figura 5 - METH3 / METH4 [Inmon '02]

Na Figura 5 estão esquematizadas as principais tarefas a desenvolver para a especificação e desenvolvimento das ferramentas de interação com o *data warehouse*. A tarefa DEPT1 visa, essencialmente, responder às necessidades da organização em termos de relatórios predefinidos, seguindo o mesmo processo de desenvolvimento que uma aplicação convencional descrita anteriormente por Inmon na fase METH1.

As tarefas IND1 a IND6 estão especialmente orientadas para o desenvolvimento do ambiente de processamento analítico de dados. O objectivo principal da tarefa IND1 é determinar quais os dados contidos no *data warehouse* que podem ser utilizados nesse tipo de processamento. Na tarefa IND2 são desenvolvidos os programas necessários para extrair os dados escolhidos durante a tarefa anterior, enquanto que a tarefa IND3 tratará dos dados obtidos e possivelmente recombiná-los-á com outros dados provenientes doutros repositórios. Por fim, na tarefa IND4, executa-se a análise dos dados para se determinar se estes respondem às necessidades do

utilizador. As duas tarefas finais visam responder às questões colocadas e à determinação se esta análise deverá ou não ser institucionalizada.

4.2 A Perspectiva de Ralph Kimball

Ralph Kimball desenvolveu uma metodologia para o desenvolvimento de projectos de SDW designada por Ciclo de Vida Dimensional do Negócio [Kimball et al. '98], que é analisada segundo três vertentes: a vertente que define o projecto do SDW, a vertente de gestão do projecto e a vertente que explicita passo a passo quais as tarefas a desempenhar de forma a que consigamos implementar correctamente o *data warehouse*.

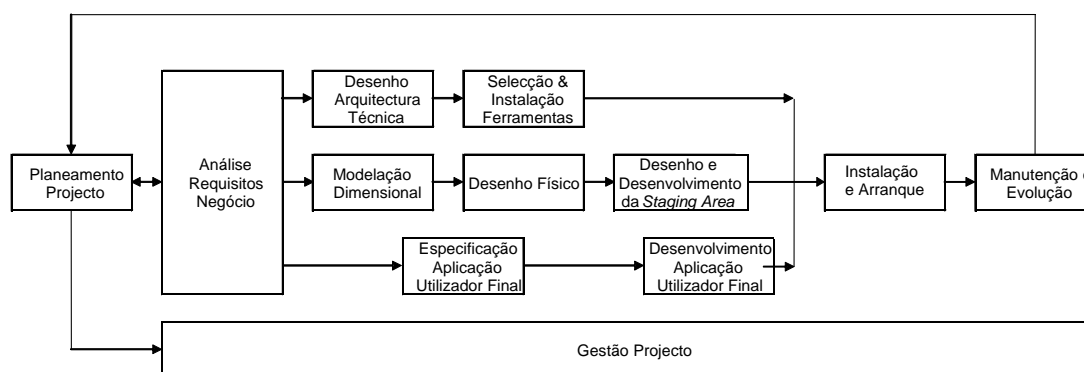


Figura 6 - Ciclo de Vida Dimensional do Negócio [Kimball et al. '98]

4.2.1 Um Plano para o Projecto

Como podemos ver na Figura 6, onde está esquematizada a metodologia apresentada por Kimball, o primeiro passo para o desenvolvimento de um SDW é a criação de um plano de projecto. A primeira tarefa a ser executada no planeamento do projecto é avaliar até que ponto a organização está preparada para embarcar num projecto desta natureza e dimensão. Para isso, Kimball propõe

o *Readiness Litmus Test* [Kimball et al. '98] que se centra, fundamentalmente, em cinco tópicos essenciais ao desenvolvimento do SDW, são eles: o patrocinador do negócio², a motivação do negócio, a parceria estabelecida entre o departamento de informática e a gerência da empresa, a cultura analítica da empresa e, por fim, a exequibilidade técnica do projecto.

O patrocinador do negócio deverá possuir algumas características essenciais para que o projecto siga o percurso mais conveniente, são elas, ter uma visão concreta sobre o impacte que um SDW terá no futuro da empresa, apoiar incondicionalmente o projecto em causa, ser um líder natural e respeitado na organização e possuir, de preferência, um historial de sucesso na empresa. Estas características conferir-lhe-ão os conhecimentos e as ligações necessárias para que seja capaz de resolver da forma mais conveniente (e expedita) as situações delicadas que eventualmente poderão ocorrer no desenvolvimento do SDW. Como forma de minimizar o risco de perda de um patrocinador, e consequente abandono do projecto, deve-se pensar sempre em obter mais do que um bom patrocinador do negócio.

Outro dos tópicos em análise é a avaliação da motivação do negócio em embarcar num projecto desta natureza. Existem diferentes motivações e a sua importância poderá contribuir decisivamente para o sucesso de implementação de um SDW. Uma forte motivação seria a existência de um mercado em permanente mudança e extremamente competitivo, o que forçaria a organização a focar-se principalmente no cliente. Outra forte motivação seria a integração dos sistemas de informação de empresas que tenham sido adquiridos recentemente pela organização. No entanto, uma fraca motivação por parte da organização em empenhar-se neste projecto aumentaria o risco de insucesso do SDW. Como exemplos, temos organizações que estão em processos de *downsizing*, ou em vias de serem adquiridas hostilmente, que não reúnem desta forma as condições ideais para apoiar o projecto.

Um projecto de *data warehousing* tem maiores probabilidades de ser um êxito quando é desenvolvido em parceria entre o departamento de informática e a gerência. O desenvolvimento autónomo de um SDW, por qualquer uma das partes, aumentaria o risco de insucesso do projecto,

² Gestor com poderes executivos dentro da organização que define usualmente os objectivos e prazos de implementação e que garante que o projecto tem os recursos, quer financeiros quer humanos, necessários à sua implementação.

sendo assim altamente desaconselhado. Ambas as partes devem colaborar e dividir responsabilidades quando decidem iniciar um projecto desta natureza.

Outro factor de grande importância, que deve ser analisado com atenção, é a cultura da empresa relativamente às tomadas de decisão. Se a organização toma decisões quanto ao seu futuro baseado no instinto dos decisores e não nos factos existentes, então um SDW não lhe trará grandes benefícios. No entanto, se essa cultura for baseada em factos, então um SDW contribuirá significativamente para apoiar os decisores nas suas funções quotidianas.

A exequibilidade do projecto é o último tópico abordado no *Readiness Litmus Test*. É necessário avaliar até que ponto está a organização preparada tecnologicamente para um projecto de *data warehousing*. Este factor é importante pois determina se a organização tem ou não uma estrutura tecnológica robusta ou se vai ser necessário um investimento adicional na reconversão tecnológica da empresa. Outra questão avaliada também neste tópico é relativa à qualidade dos dados existentes nos sistemas fontes e da disponibilidade de recursos experientes para a participação neste projecto. Kimball não atribui igual importância aos diversos factores analisados. O factor mais influente e crítico, no desenvolvimento do projecto, é sem dúvida o factor relativo ao patrocinador do negócio. Comparativamente, Kimball atribui cerca de 60% de importância a este elemento e cerca de 15% à motivação do negócio e à exequibilidade técnica, ficando cerca de 10% para a parceria entre o departamento de informática e a gerência, e para a cultura analítica da empresa.

A segunda tarefa a executar no planeamento do projecto do SDW é determinar qual o âmbito do projecto e a justificação dos requisitos de forma a determinar qual o financiamento necessário. Esta tarefa deve ser desempenhada em conjunto pelo departamento de informática e pela gerência, sempre com a consciência que o âmbito deve ser definido tendo por base os requisitos do negócio identificados e não as questões temporais. O âmbito do projecto deve ser assim definido tendo, também, por base o impacte que este vai ter na organização. Devemos começar por identificar apenas um requisito de negócio referente a um só processo, mas que este tenha impacte na organização. É também necessário identificar quantos utilizadores vão ser abrangidos por esta primeira iteração do *data warehouse*, partindo sempre de um número limitado de forma a

que o projecto não se torne de difícil gestão. Por último, deve-se estabelecer alguns critérios de sucesso após implementação do SDW. Desta forma a gerência terá uma ideia concreta acerca dos benefícios alcançados com a sua implementação.

Após a definição do âmbito do projecto, deve o departamento de informática, sempre em parceria com a gerência da empresa, avaliar a justificação dos requisitos através da análise do custo e do investimento necessário à realização desta iteração, assim como o retorno e benefícios alcançados com a mesma. No cálculo dos custos e do investimento necessário para a realização deste projecto devemos incluir factores como despesas efectuadas com a aquisição de *hardware* e *software*. É necessário também quantificar as despesas que irão ser efectuadas com a manutenção do *hardware* e *software* assim como as despesas projectadas para um crescimento futuro do SDW. Os custos relativos aos recursos humanos devem também ser contabilizados, sejam eles referentes a recursos internos ou externos, assim como as despesas que se terão com a formação dos membros do projecto e dos utilizadores do SDW. Por fim, quando o SDW estiver em funcionamento é necessário também contabilizar os custos despendidos com o contínuo suporte do sistema.

A avaliação dos benefícios e do retorno do investimento obtido com a conclusão deste projecto deve ser realizada com algum detalhe. Esta avaliação é bem mais difícil de executar que a determinação dos custos. Para ser possível uma avaliação realista do retorno do investimento efectuado, é necessário avaliar o rendimento, ou o custo existente antes da existência do SDW, e projectar o rendimento ou custo obtido com o uso de um sistema, que permitirá obter dados mais rapidamente e com maior precisão. Operações como *marketing*, *cross-selling*, custos com matérias primas, saíram com certeza beneficiadas com a utilização de um SDW.

Finalmente, deve ser elaborado um documento onde são descritas todas as questões abordadas nesta tarefa, identificando principalmente qual o âmbito desta iteração, quais os tópicos que ficaram fora do âmbito desta iteração e que serão abordados no futuro, quais os critérios de sucesso e quais as medidas a tomar para os possíveis riscos que poderão emergir nesta iteração, juntamente com a justificação dos requisitos onde estão identificados o custo, o investimento, assim como o retorno e benefícios alcançados com a conclusão deste processo.

Uma vez obtida a aprovação e o financiamento necessário à execução do projecto do SDW, deve ser elaborado um plano mais detalhado, onde se deve nomear o projecto e identificar quais os elementos que vão compor cada um dos grupos de trabalho bem como as tarefas e sub-tarefas a desempenhar de acordo com o calendário preestabelecido.

Existem assim vários grupos de trabalho necessários à execução das mais variadas tarefas durante o decorrer de todo o projecto. Como já foi referido, para um projecto desta natureza evoluir sem grandes atritos e com uma baixa taxa de risco é necessário que existam um ou mais patrocinadores do negócio. Existe assim um grupo composto pelos patrocinadores de negócio com a tarefa de supervisionar o projecto, fornecer financiamento necessário à execução do projecto, cativar futuros aderentes ao projecto e desbloquear problemas que possam surgir no decorrer da implementação do projecto.

O grupo que está hierarquicamente a seguir a este é composto pelo gestor do projecto e pelo grupo que representa a gerência da empresa. Este grupo tem como principal tarefa a gestão e monitorização do projecto, mais especificamente o controlo das tarefas a realizar e dos recursos a alocar. O perfil do gestor do projecto é extremamente importante para a conclusão com sucesso do projecto. Este deve ter um bom conhecimento sobre tecnologia e desenvolvimento de sistemas, uma grande capacidade comunicativa e boas ligações com a gerência da organização. Este cargo exige que seja desempenhado a tempo inteiro já que exige uma total dedicação.

A equipa que tem como tarefa o desenho e desenvolvimento do SDW está sobre a coordenação do gestor do projecto e é composta por elementos com diferentes perfis, definidos de acordo com as necessidades de cada um dos projectos. De salientar que os elementos poderão acumular mais do que uma tarefa, variando essa situação conforme a dimensão do projecto em causa. Os perfis necessários ao desenho e desenvolvimento do projecto são, normalmente, os seguintes:

- O **analista do negócio**, que é normalmente um consultor externo especializado na área de negócio em que se posiciona a organização. Tem como responsabilidade a análise e

projectão da evolução do negócio com vista a uma melhor adaptação ao futuro da organização em questão.

- O **analista de dados** que tem como principal tarefa a interpretação dos dados presentes nas diversas fontes com o intuito de no futuro poder desenvolver o modelo dimensional de dados.
- O **administrador das bases de dados** do SDW que é o responsável pela transformação do modelo dimensional em estruturas físicas de armazenamento de dados. Tem também como principal responsabilidade planear esquemas de agregação, particionamento e criação de índices. Uma vez implementado o *data warehouse* o DBA fica responsável pela manutenção diária da base de dados, garantindo a sua disponibilidade, integridade e níveis de performance pré-determinados.
- O **projectista do sistema da área de retenção**, que é responsável pelo desenho e desenvolvimento das ferramentas ETL dos dados para o *data warehouse*.
- Os **programadores** desenvolvem as aplicações que os utilizadores finais utilizarão para interagir com o *data warehouse*.
- Os **formadores** em *data warehousing* que organizarão e transmitirão o conhecimento necessário sobre os SDW aos utilizadores finais, para que estes consigam, através do uso correcto das aplicações finais, aceder ao *data warehouse* e, como tal, analisar os dados neles existentes.

Existem também algumas equipas especiais que poderão aparecer conforme a necessidade do projecto e que são responsáveis por tarefas específicas, como por exemplo, a qualidade, a segurança ou o suporte do sistema e dos seus dados.

Uma vez definidas as equipas que vão integrar o desenvolvimento do SDW é necessário desenvolver o plano do projecto, identificando as tarefas associadas, os recursos envolvidos na tarefa, as datas de início e conclusão do projecto, a percentagem de trabalho realizado, os dias previstos para a execução da tarefa, os dias despendidos na execução da tarefa e os dias de atraso do projecto.

Como conselho, Kimball sugere que as previsões temporais necessárias à conclusão de todas as tarefas sejam inflacionadas de forma a precaver um pouco o futuro do projecto. Deve ser prestada especial atenção à estimativa de tempo necessária à execução da área de retenção. É quase impossível prever exactamente quanto tempo será necessário para a conclusão de tal tarefa, sendo esta por vezes a razão do atraso de muitos projectos de *data warehousing*. Como tal, é recomendado que: uma vez previsto o tempo necessário para o processo referente à área de retenção, deve-se quadruplicá-lo.

Avaliação da Organização	<ul style="list-style-type: none">• <i>Readiness Litmus Test</i>
Planeamento do Projecto	<ul style="list-style-type: none">• Definição do âmbito.• Orçamentação (Planeamento dos Custos e Benefícios).
Desenvolvimento do Plano	<ul style="list-style-type: none">• Definição do Gestor do Projecto.• Selecção/Contratação dos elementos da equipa de desenvolvimento.• Criação dos grupos de trabalho.• Atribuição de tarefas e responsabilidades aos grupos de trabalho.• Calendarização das tarefas.

Tabela 2 – Resumo das tarefas a executar na elaboração do projecto segundo Kimball

4.2.2 Gestão

Uma vez elaborado o plano detalhado do projecto de *data warehousing* é necessário despende o tempo adequado para a monitorização e gestão do projecto. As técnicas sugeridas por Kimball para uma gestão eficaz de um projecto de *data warehousing* começam pela realização de uma reunião com todos os elementos participantes no projecto de *data warehousing*. Esta reunião tem como objectivo sintonizar todos os participantes no projecto para um mesmo objectivo. Deve ser iniciada pelo patrocinador do negócio, identificando os objectivos do projecto e o seu âmbito, e seguidamente conduzida pelo gestor de projecto. Nesta reunião deve ser feito um resumo do projecto, a identificação das equipas de trabalho e suas responsabilidades, e a apresentação do plano do projecto.

Kimball sugere que se faça, como técnica eficaz para a boa gestão do projecto de *data warehousing*, uma monitorização constante do plano de projecto, uma vez que os diferentes

recursos e tarefas deverão convergir para um ponto em comum. Só assim é que se terão as condições necessárias para se atingir o sucesso do projecto. Como forma de melhor monitorizar o projecto, deverão os membros da equipa principal reunir todas as semanas, para se fazer um ponto da situação e relatar eventuais problemas ocorridos durante o desenvolvimento, bem como comunicar as soluções encontradas. Deverão também ser abordadas as questões por resolver, com o intuito de procurar uma solução. O resultado destas reuniões semanais deverá ser descrito num relatório, que deverá ser apresentado periodicamente a todos os elementos da equipa principal e também ao patrocinador do negócio. Este relatório não deverá ser muito pormenorizado. Porém, deverá relatar o trabalho até então desenvolvido assim como o trabalho planeado para o período seguinte e os problemas ainda por resolver.

Uma das tarefas mais ingratas que deverá ter uma atenção especial por parte do gestor do projecto é a documentação do próprio projecto. O gestor do projecto não deverá cair na tentação de adiar a gestão da documentação do projecto sob pena de esta se acumular sucessivamente até atingir níveis indesejáveis. A documentação deverá reflectir a evolução do projecto, seja este favorável ou não ao planeamento original. Só desta forma se consegue diagnosticar antecipadamente alguns dos problemas mais comuns como o atraso na execução de uma tarefa ou o esgotamento do orçamento atribuído a uma tarefa – que, como se sabe, terão um impacte significativo no resto do projecto.

O Gestor do Projecto tem também a árdua tarefa de gerir o âmbito do projecto. À medida que o projecto se desenvolve poderão surgir questões que não foram previamente identificadas, tais como um sistema fonte não identificado, má qualidade dos dados, ou requisitos adicionais de última hora. Kimball sugere, caso uma destas situações ocorra, que o gestor do projecto tome uma de três decisões:

- negar a inclusão de tais requisitos no projecto;
- ajustar o âmbito do projecto, mas de forma a que o esforço inicial se mantenha;
- adicionar os novos requisitos, aumentar o âmbito do projecto e recorrer a um aumento de financiamento e tempo para a conclusão do projecto.

À medida que o projecto decorre surgem, naturalmente algumas mudanças ou problemas que poderão ter impacte a nível do custo ou do tempo alocado na implementação do SDW. É necessário que sejam registadas todas estas ocorrências através de um registo próprio no qual estejam identificadas as questões, data da ocorrência, data de resolução, o estado, o tipo de prioridade e quem as resolveu. Caso estas ocorrências impliquem mudanças sobre o que foi inicialmente projectado, estas deverão ser também registadas e avaliado o impacte a nível de esforço e custo para o projecto do SDW. Esta informação deverá ser depois comunicada ao patrocinador de negócio do projecto.

Finalmente, o gestor do projecto deverá estabelecer um plano de comunicação periódico com o objectivo de melhor gerir as expectativas de todos os que estão envolvidos neste projecto. À partida estão identificados três grupos que o gestor deverá prestar atenção: a equipa principal de desenvolvimento do SDW, os patrocinadores de negócio e os utilizadores finais do sistema. No entanto, poderá ser necessário ocasionalmente apresentar a evolução do projecto à gerência, a elementos do departamento de informática não envolvidos neste projecto e, por fim, à organização em geral.

Reunião Inicial	<ul style="list-style-type: none">• Iniciada pelo Patrocinador de Negócio.• Desenvolvida pelo Gestor do Projecto.
Reuniões Periódicas	<ul style="list-style-type: none">• Elementos da Equipa.• Tema abordados:<ul style="list-style-type: none">◦ Tarefas concluídas.◦ Tarefas atrasadas.◦ Problemas pendentes.• Resumo da reunião enviado para o Patrocinador de Negócio.
Gestão do Âmbito	<ul style="list-style-type: none">• Gestor do Projecto.• Patrocinador do Negócio.• Reajustamento do plano do projecto, caso haja alterações aos requisitos, isto é, quando sucedam alterações ao âmbito do projecto que comprometam o plano inicial.
Plano de Comunicação	<ul style="list-style-type: none">• Elaboração de relatórios de estado.• Distribuição pelos membros da equipa.• Distribuição pelos Patrocinadores de Negócio.• Distribuição pelos Utilizadores Finais.

Tabela 3 – Resumo das tarefas a executar na gestão de um projecto segundo Kimball

4.2.3 Implementação

Relativamente à implementação do SDW, Kimball define que este processo pode ser executado recorrendo a três linhas de acção de desenvolvimento paralelo. A saber: a linha de acção tecnológica, a linha de acção dos dados e a linha de acção das aplicações. Esta definição permite uma redução do tempo de conclusão do projecto uma vez que as tarefas podem ser executadas em paralelo, especialmente aquelas que requerem recursos humanos com perfis diferentes.

Linha de Acção Tecnológica

O objectivo do conjunto de tarefas que integram esta linha de acção é o estabelecimento da arquitectura técnica que servirá de suporte ao SDW, que considerará a análise das infra-estruturas existentes, a análise dos requisitos e a implementação da solução atingida. Kimball identifica oito passos para a criação da arquitectura técnica, começando pela constituição de uma equipa dedicada à construção da arquitectura técnica. Esta equipa deverá recolher todos os requisitos técnicos e documentá-los apropriadamente. Tendo por base esses requisitos, deverá ser elaborada uma arquitectura modelo de alto nível que, posteriormente, terá de ser especificada detalhadamente. Após terem sido identificados todos os subsistemas que compõem a arquitectura técnica, é necessária a elaboração de um plano de implementação organizado por fases. Finalmente, a equipa deverá documentar todos os planos detalhadamente para os técnicos poderem seguir as instruções na implementação do plano. A finalizar, a equipa dedicada à arquitectura técnica deverá proceder ao esclarecimento e formação de todos os intervenientes no projecto do SDW.

A tarefa seguinte, desenvolve-se em torno da selecção e instalação dos produtos necessários à implementação da arquitectura técnica que servirá de suporte ao SDW, envolvendo, nomeadamente, a aquisição de *hardware* ou *software*. Nesta tarefa deve ser prestada atenção ao processo de aquisição de bens na organização, pois tal processo pode ser periódico e poderá influenciar temporalmente o processo de aquisição de *software* e *hardware*.

Uma vez elaborado o documento com a arquitectura técnica, é necessário criar uma matriz de avaliação segundo diferentes critérios e com diferentes pesos para poder ser escolhido o produto que mais se adequa às necessidades do projecto. Existem em todos os produtos prós e contras, e estes devem ser convenientemente analisados e ponderados de acordo com as necessidades do projecto.

Kimball aconselha também que se efectue uma análise de mercado para que se possa entender melhor o que os fornecedores têm a oferecer e também como forma de melhor se perceber aquilo que pode ser pedido aos fornecedores, relativamente a servidores, sistemas de gestão de bases de dados ou sistemas de gestão de cópias de segurança. Após uma primeira análise dos produtos ao dispor, dever-se-á proceder à sua avaliação detalhada. Finalmente, uma vez escolhido o produto, este deve ser instalado para testes e negociada a sua eventual aquisição.

Linha de Acção dos Dados

Uma vez definido o processo de negócio a ser abordado nesta iteração do sistema de *data warehousing*, é agora necessário uma análise mais detalhada dos dados gerados por este processo. Mais tarde, será elaborado o esquema dimensional que suportará o processo de negócio em causa. O esquema dimensional deverá ser apresentado e validado quer pelo restante grupo do projecto de *data warehousing* quer pelo grupo de utilizadores que acompanha este processo. Todo este processo de modelação dimensional deve estar documentado ao pormenor para uma eventual consulta, esclarecimento ou justificação futura.

Concluído o modelo dimensional, passa-se à fase de tradução desse modelo para o correspondente esquema físico. Durante este processo deverão ser contempladas todas as acções que poderão melhorar o desempenho do SDW. Complementarmente à tradução do modelo dimensional deverão ser equacionadas algumas estratégias de optimização base como a introdução de elementos pré-calculados, a criação de índices, e a eventual desnormalização de algumas estruturas de dados.

A última tarefa a desempenhar nesta linha de acção é o desenho e a construção da área de retenção. Este é um processo em tudo semelhante ao da arquitectura tecnológica. Começa-se por

desenhar um modelo de alto nível, sendo posteriormente realizada a especificação detalhada de cada um dos componentes envolvidos, neste caso de cada uma das tabelas que integrem o esquema físico do *data warehouse*. A finalizar, serão especificadas as diversas ferramentas para a extracção, transformação e integração de dados no *data warehouse* da organização. Kimball aconselha o uso de ferramentas comerciais para a execução dessas tarefas. Dever-se-á ter em conta que, o uso de ferramentas comerciais poderá trazer encargos adicionais na compra de *software*, mas poderá minimizar o tempo de conclusão do projecto comparativamente ao desenvolvimento interno da ferramenta.

Linha de Acção das Aplicações

O conjunto de tarefas envolvidas nesta linha de acção permitem que se comece a ver o que o futuro do SDW vai ser. É nesta fase que se inicia a integração de todo o SDW através do desenvolvimento de ferramentas que serão utilizadas por utilizadores muito especiais (os agentes de decisão das organizações) que tirarão partido do investimento efectuado na arquitectura tecnológica e nos dados.

Partindo do mesmo ponto que as linhas de acção abordadas anteriormente, isto é, a partir do documento com os requisitos do negócio, é agora necessário especificar o conjunto das aplicações analíticas, sejam estas ferramentas de descoberta de conhecimento, ferramentas para a elaboração de relatórios ou ambientes para o lançamento de questões esporádicas. Estas aplicações deverão ter um interface comum, ou pelo menos semelhante, e deverão permitir uma interacção fácil e amigável entre o utilizador e o sistema. A implementação destas aplicações deverá ser feita assim que a especificação estiver concluída e logo que o desenho da base de dados estiver concluído. Nestas operações deverão ser respeitados todos os *standards* estabelecidos na especificação da apresentação de resultados, cálculo de atributos derivados e convenções de nomes.

Integração das Três Linhas de Acção

O processo de convergência destas três linhas de acção tem o nome de instalação, processo esse que dada a sua dificuldade deve ser planeado com muito cuidado. Associado a este processo, existe também a componente de formação e suporte que terá muita importância no futuro do SDW, pois irá permitir a integração dos utilizadores no projecto.

Finalmente, existe a última fase desta iteração que é dedicada à manutenção e ao planeamento do crescimento do SDW. Como um projecto destes, caso seja bem sucedido, “nunca” tem um fim, é necessário providenciar os mecanismos adequados para a realização de processos de manutenção permanentes sobre *data warehouse*. Assim, conseguir-se-á responder aos requisitos dos utilizadores, sejam estes novos requisitos ou simples ajustamentos de requisitos prévios. Se o SDW for bem sucedido, este irá atrair, com certeza novos utilizadores e novas funcionalidades.

4.3 Outras Perspectivas

4.3.1 Sid Adelman

Sid Adelman e Larissa Moss editaram um livro [Adelman et al. '00] sobre gestão de projectos de *data warehousing*, livro esse que, como o título indica, foca com especial atenção uma parte do projecto que não tem sido abordada com a profundidade necessária pelos outros autores da área e que tem bastante influência no destino de um projecto de *data warehousing* : a gestão do projecto. Adelman [Adelman et al. '00, Adelman et al. '98] apresenta várias questões a ter especial atenção, ora indicando meramente a sua possível ocorrência, ora sugerindo soluções para os mais variados problemas.

Após a análise das diferentes propostas preconizadas por Inmon [Inmon '02] e Kimball [Kimball et al. '98], existem ainda algumas falhas relativas a componentes que ainda não foram abordadas com a especificidade recomendada. Adelman dá uma visão menos técnica do projecto de *data warehousing* focando problemas como a definição de objectivos de curto e longo prazo de um *data*

warehouse e os passos para a sua obtenção. Define também que, num projecto desta natureza, devem ser descritas quais as medidas de sucesso utilizadas, de forma a avaliar o comportamento do SDW, e quais os factores críticos a ter em atenção, enunciando também como quantificar o sucesso do projecto. Dada a alta taxa de insucesso de implementação de SDW, Adelman presta especial atenção à identificação do tipo de falhas mais comuns neste projectos [Adelman '01, Adelman et al. '00], quais os tipos de riscos que afectam o projecto e quais as medidas a adoptar para os minorar. Este considera que um dos factores com maior importância para o sucesso de um projecto de um SDW são os seus utilizadores, daí que sejam enunciados os diferentes tipos de utilizadores existentes neste ambiente e sugeridas formas de comunicação para a obtenção do apoio e compreensão dos mesmos. Outro factor importante, referido por ele, é a gestão do orçamento atribuído ao projecto e quais as acções que devem ser preconizadas pelo gestor de forma a controlar os gastos e como justificá-los perante a gerência.

A selecção do *software* que irá ser utilizado – ferramentas ETI, de modelação, ou de consulta de dados – tem também grande influência quer sobre o orçamento do próprio projecto quer sobre as suas limitações temporais. Adelman descreve qual o propósito de cada uma dessas ferramentas, relevando a sua importância e algumas das técnicas utilizadas na sua selecção. No seu livro são também abordadas questões relacionadas com a modelação dos dados, qualidade dos dados e cultura organizacional, como factores que têm alguma influência sobre o projecto de um SDW.

A metodologia descrita por Adelman tem por base passos e não fases. Este considera que a palavra fase tem uma conotação ligada com o desenvolvimento através da perspectiva *waterfall* convencional e, como tal, não a adopta. Define também que cada passo pode atravessar mais que uma fase tradicional e que pode ser repetitivo.

O primeiro passo é denominado como acordo de projecto. Neste é elaborado um documento com diversas secções onde deverão estar identificados, por exemplo:

- os objectivos a atingir com o SDW, os requisitos de negócio e respectiva solução;
- a análise custo-benefício e o âmbito do projecto (e o que não pertence ao âmbito do projecto);

- o modelo de dados de alto nível para o tema escolhido;
- as infra-estruturas e os processos de análise da qualidade dos dados das fontes de informação;
- a quantificação das fontes de dados;
- a atribuição de tarefas e responsabilidades dos intervenientes e dos elementos da equipa do projecto;
- a disponibilidade dos recursos e seus requisitos de segurança;
- as ferramentas de acesso aos dados.

No passo seguinte deve-se rever e analisar qual a plataforma tecnológica a adoptar, nomeadamente a escolha do CPU, a estratégia de armazenamento dos dados, o sistema de gestão de bases de dados relacional, as ferramentas ETI, as ferramentas de criação de consultas e relatórios, entre outras. Paralelamente à actividade anterior, outros membros da equipa do projecto poderão concentrar-se nos processos de ETI do SDW.

No decorrer de cada um desses passos, se houver alguma alteração ao dito acordo de projecto, este deve ser revisto e alterado. É neste passo que se conclui o modelo lógico de dados e é feita a análise rigorosa das fontes de dados. Se houver descoberta de novos atributos estes são comunicados à equipa de modelação que tratará de actualizar o modelo lógico de dados. De seguida, as ferramentas ETI são analisadas e desenvolvidas se necessário, e mediante o destino dos dados, seja ele um *Operational Data Store* (ODS) ou um *Enterprise Data Warehouse* é feito o desenho da base de dados.

Ao mesmo tempo que o desenvolvimento das ferramentas ETI, pode ocorrer a especificação e o desenvolvimento das ferramentas de consulta e de elaboração de relatórios. Neste passo são também desenvolvidas as actividades relacionadas com o suporte do SDW, com a formação dos utilizadores finais, com a elaboração de testes ao sistema e com a preparação para a sua implementação. Para finalizar, restam ainda dois passos: a implementação do sistema e a pós-produção. A implementação do sistema só deve ser feita após se terem realizado testes durante um ou dois meses com alguns utilizadores seleccionados. A implementação do sistema passa por disponibilizar as ferramentas de acesso ao *data warehouse*, pela sua alimentação total, pela

formação dos restantes utilizadores e pela finalização do projecto. A pós-produção é uma tarefa constante de manutenção e de preparação para a próxima iteração.

4.3.2 Michael Gonzales

Michael Gonzales editou um livro [Gonzales '03] sobre a perspectiva da IBM no mundo dos *data warehouses*. A análise efectuada nessa obra é exclusivamente sobre o capítulo dedicado ao planeamento de iterações na implementação de um SDW. A visão do autor sobre este planeamento é independente das plataformas a seleccionar o que torna a sua perspectiva de aplicação genérica. Gonzales analisa a construção adequada de um projecto de um SDW, identificando quais as principais áreas que o plano deverá abordar. De referir, a parte estratégica do desenvolvimento, a arquitectura de dados, a arquitectura técnica e finalmente a implementação e manutenção do sistema.

Como o desenvolvimento de um SDW não deve ser feito de um passo só, dado o risco que a dimensão do projecto pode acarretar, o autor defende que a construção seja feita de forma iterativa, com objectivos de curto-prazo bem definidos, mas tendo sempre em atenção o planeamento de longo-prazo do SDW. A base para a criação de um plano de construção para um SDW é um documento que está em constante mudança que descreve e define todas as componentes envolvidas no projecto. Este documento é como um mapa, é uma espécie de guia para uma implementação com sucesso. A nível estratégico, o documento deverá conter informação acerca dos objectivos e constrangimentos da arquitectura, da vista conceptual do *data warehouse* e dos *standards* adoptados na implementação do projecto. Relativamente à parte técnica, o documento deverá definir, essencialmente, quais os seus objectivos base, a arquitectura tecnológica, seus constrangimentos, e componentes que deverão ter especial atenção. A componente de dados do sistema deve ser de igual modo abordada. Finalmente, devem ser definidas as (melhores) estratégias a seguir para se alcançar com sucesso o final do projecto. Como recomendação, Gonzales diz que o plano de desenvolvimento deve no seu início conter diagramas de alto nível, algumas directivas gerais e definições genéricas acerca do projecto. À medida que o projecto evolui o conteúdo da documentação deverá também ser, naturalmente, aprofundado.

Uma vez que o processo de desenvolvimento de um SDW é iterativo é preciso prestar muita atenção à sua primeira iteração, pois esta é bastante crítica e ao mesmo tempo a mais longa de todas as iterações. Tal deve-se à necessidade de se realizar os processos de identificação e definição das estratégias de desenvolvimento de todo o SDW, que envolvem questões um pouco delicadas como a definição da arquitectura de dados e da arquitectura tecnológica, a gestão dos metadados locais e empresariais e a implementação dos diversos métodos de acesso aos sistemas de dados da empresa. É necessário também fazer-se a identificação das ferramentas necessárias para a realização dos processos de extracção, transformação e integração de dados, gestão dos metadados e acesso aos dados. O tão necessário processo de formação dos futuros utilizadores do SDW termina esta fase do projecto.

No planeamento da primeira iteração o gestor do projecto deve alinhar o SDW com a estratégia da organização. O SDW deverá ser desenvolvido prioritariamente no sentido de resolver e tratar as questões estratégicas relacionadas com as actividades do negócio das empresas. Nesse sentido, foram definidos quatro princípios que devem ser seguidos durante o desenvolvimento do SDW que são, nomeadamente, a identificação do patrocinador, a focalização do desenvolvimento nos requisitos estratégicos, o saber que são requisitos do negócio que determinam as várias iterações do SDW e o reconhecimento que este é um bem em permanente desenvolvimento.

A tarefa seguinte será tratar da adopção de um método formal de avaliação para a organização, de forma a identificar eventuais problemas que possam ocorrer. O gestor do projecto deve elaborar um documento que avalie se a empresa está ou não preparada para o esforço de implementação de um SDW. Para tal, esse documento deve focar, se possível, os seguintes tópicos:

- Identificar sucessos/insucessos de antigos projectos de SDW.
- Identificar possíveis áreas problemáticas e deficiências a resolver.
- Perceber as capacidades tecnológicas da empresa.
- Compreender a cultura empresarial.
- Identificar eventuais deficiências organizacionais ou de recursos.

É também necessário promover e realizar uma avaliação dos recursos humanos envolvidos no projecto do SDW. De referir, que em projectos de pequena dimensão um mesmo elemento da equipa de desenvolvimento pode acumular várias funções – as funções mais típicas existentes num projecto de um SDW estão apresentadas na Tabela 4. De salientar que uma equipa típica de um projecto SDW costuma ter alguns elementos a tempo inteiro e outros a tempo parcial, distribuídos normalmente da seguinte maneira:

- **Tempo Inteiro:** gestor projecto, arquitecto de sistemas dados, programador extracção de dados e programador acesso aos dados.
- **Tempo Parcial:** administrador de base de dados, analista requisitos negócio e especialista de Negócio.

É vivamente aconselhado que em futuras iterações se faça a distribuição dos membros da equipa do projecto mais experientes por novas equipas e juntar novos membros, de forma a minimizar o tempo de aprendizagem dos novos elementos e o tempo de entrega das iterações seguintes do projecto.

Patrocinador do negócio	Gestor executivo com poderes para assegurar os recursos necessários ao projecto.
Gestor do projecto	Responsável pelo planeamento e desenvolvimento do SDW.
Analista requisitos negócio	Consultor que servirá de intermediário entre a equipa de desenvolvimento e os utilizadores finais com vista ao esclarecimento de dúvidas.
Especialista do negócio	Consultor especializado no Negócio da Organização.
Arquitecto de dados	Responsável pelo desenho lógico do DW.
Arquitecto técnico	Responsável pelo desenho e implementação das infra-estruturas técnicas de suporte ao DW.
Administrador metadados	Responsável pelo armazenamento de todos os procedimentos e metadados do DW.
Administrador de base de dados	Responsável pelas bases de dados operacionais.
Administrador do SDW	Responsável pela implementação do desenho lógico do DW e administração do DW.
Programador de processos extracção de dados	Define, desenha e implementa estratégias de acesso aos dados dos sistemas fontes.
Programador de acessos aos dados	Tem como principais tarefas o desenho e desenvolvimento das ferramentas de acesso ao DW.

Tabela 4 – Funções típicas existentes num projecto de DW

A finalizar, é necessário escolher o tema a implementar no SDW. Gonzales utiliza um método formal [Gonzales '03] para determinar o tema mais importante e com maior impacte para a organização. Esse método foi designado por Matriz DIF. Esta matriz tem como principal objectivo analisar os requisitos a implementar, isto é quais os temas candidatos, e avaliá-los segundo três pontos essenciais: disfuncionalidade (*Dysfunctional*), impacte (*Impact*), e viabilidade (*Feasibility*). Esta avaliação é feita normalmente por duas comunidades diferentes, os utilizadores e os técnicos. Nela são calculadas as médias das avaliações de cada comunidade assim como a média total para cada ponto de cada processo. No final podem-se ainda atribuir pesos diferentes a cada um dos três pontos referidos e calcular novamente a média total pesada. A partir deste ponto é identificado o processo para a próxima iteração. Está, assim, garantida a escolha do processo mais importante para todos e para o futuro da empresa. Todavia, Gonzales não enuncia nenhuma metodologia de implementação para um SDW, mas defende que, das metodologias actualmente existentes – Kimball, Inmon, em espiral ou a IBM -, deve-se escolher a que melhor se adaptar às necessidades particulares da organização que irá acolher o SDW.

Capítulo 5

Gestão de Projectos de *Data Warehousing*

5.1 A Gestão de Projectos

Para que seja possível definir o que é a gestão de projectos e quais as técnicas e métodos associados a uma gestão de projectos eficaz, é necessário primeiro definir o termo projecto. A identificação e definição de um projecto leva a uma clarificação do termo, e do que este representa, sendo então possível definir as tarefas associadas à sua gestão de forma a atingir-se os objectivos definidos.

A *framework*³ estudada está contida no guia [PMI '00] da *Project Management Institute*, sendo já um *standard de facto* e é utilizada na gestão de projectos. A PMI é uma associação profissional sem fins lucrativos na área de gestão de projectos, organizando conferências, promovendo a estandardização da gestão de projectos, e a certificação dos gestores de projectos.

³ Conjunto de passos e tarefas a desempenhar para se atingir um objectivo.

5.1.1 Projecto

No nosso dia-a-dia as organizações efectuem trabalhos, esses trabalhos envolvem normalmente a execução de operações ou projectos. É pois necessário haver uma clarificação entre o significado de ambos os termos, uma vez que existem pontos em comum e diferenças entre ambos, originando a má aplicação dos termos, das técnicas, e consequente confusão.

As características em comum entre operações e projectos são o facto de ambos serem desempenhados por pessoas, ambos estarem limitados por um conjunto de recursos e a sua execução ser alvo de um planeamento cuidado e controlo constante. No entanto, a principal diferença é que as operações são constantes e repetitivas ao invés dos projectos, que são únicos e temporários [Wideman '95].

Os projectos são normalmente executados como um meio para atingir um objectivo de um plano estratégico da organização. Pode-se então definir um projecto como um esforço temporário para a criação de um produto ou serviço único. Diz-se que um projecto é temporário porque este tem um início determinado, assim como uma previsão para a sua conclusão. Diz-se previsão de conclusão porque um projecto chega à fase de conclusão quando este atinge os seus objectivos, ou quando se torna claro que os objectivos do projecto não serão ou não poderão ser atingidos ou que a necessidade do projecto deixou de ser válida e, como tal, o projecto será terminado. Quando se define que um projecto é único, é porque a sua execução é de alguma forma diferente dos projectos até então executados, o que lhe confere tal estatuto. O exemplo mais comum seria o da construção de um edifício. Este, é um projecto único pois a sua execução será de alguma forma diferente dos projectos de construção de outros edifícios, ou porque o proprietário e o local são diferentes, ou pelo facto dos empreiteiros e materiais utilizados serem também diferentes.

5.1.2 *Framework* Aplicada à Gestão de Projectos

A gestão de projectos é a aplicação de conhecimento, ferramentas e técnicas às actividades de um projecto com o intuito de este atingir os seus objectivos com sucesso. A gestão de projectos contempla um conjunto de áreas de conhecimento, onde são considerados um conjunto de processos. Estes processos estão classificados segundo cinco grupos de processos, o grupo de processos de Iniciação, o grupo de processos de Planeamento, o grupo de processos de Execução, o grupo de processos de Controlo e o grupo de processos de Encerramento (Figura 7).

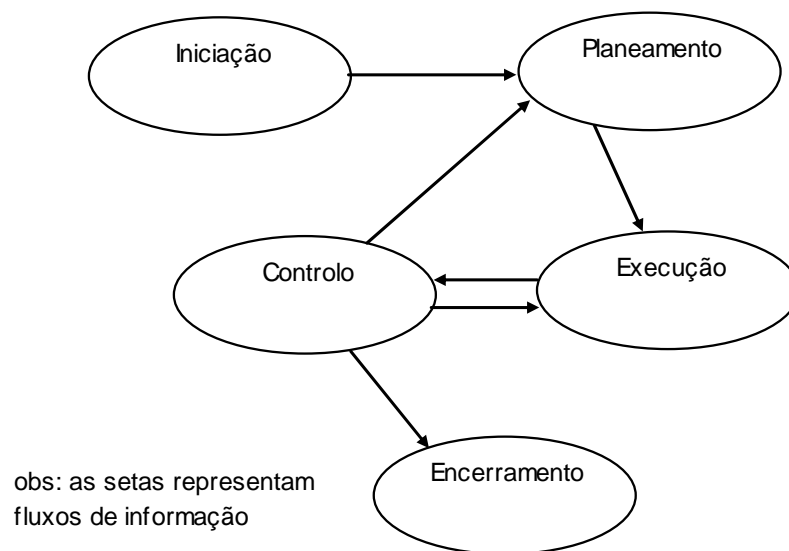


Figura 7 - Grupos de Processos [PMI '00]

Os processos de Iniciação lidam com a autorização para o início do projecto ou de uma determinada fase. Os processos de Planeamento definem (ou refinam) objectivos através da escolha de um plano de acção para a obtenção dos objectivos a que o projecto está sujeito. Os processos de Execução coordenam os recursos humanos e materiais para a execução de um plano previamente definido. Os processos de Controlo visam garantir que os objectivos são atingidos através da constante monitorização das tarefas, registando o progresso de forma a que sejam possíveis acções correctivas na eventualidade de um desvio do plano inicial. Os processos de Encerramento formalizam a aceitação da conclusão do projecto ou fase determinando o seu encerramento.

Foram identificadas pela PMI nove áreas científicas que deveriam ser do domínio de todos os gestores de projecto. À imagem de outras profissões, onde é necessário o domínio absoluto de determinadas áreas, o gestor do projecto deverá dominar todas as áreas do conhecimento que fazem parte do desenvolvimento de um projecto e todos os seus processos.

As áreas de conhecimento abordadas na gestão de projectos são a Gestão da Integração, a Gestão do Âmbito, a Gestão do Tempo, a Gestão do Custo, a Gestão da Qualidade, a Gestão dos Recursos Humanos, a Gestão da Comunicação, a Gestão do Risco e a Gestão de Compras/Subcontratação.

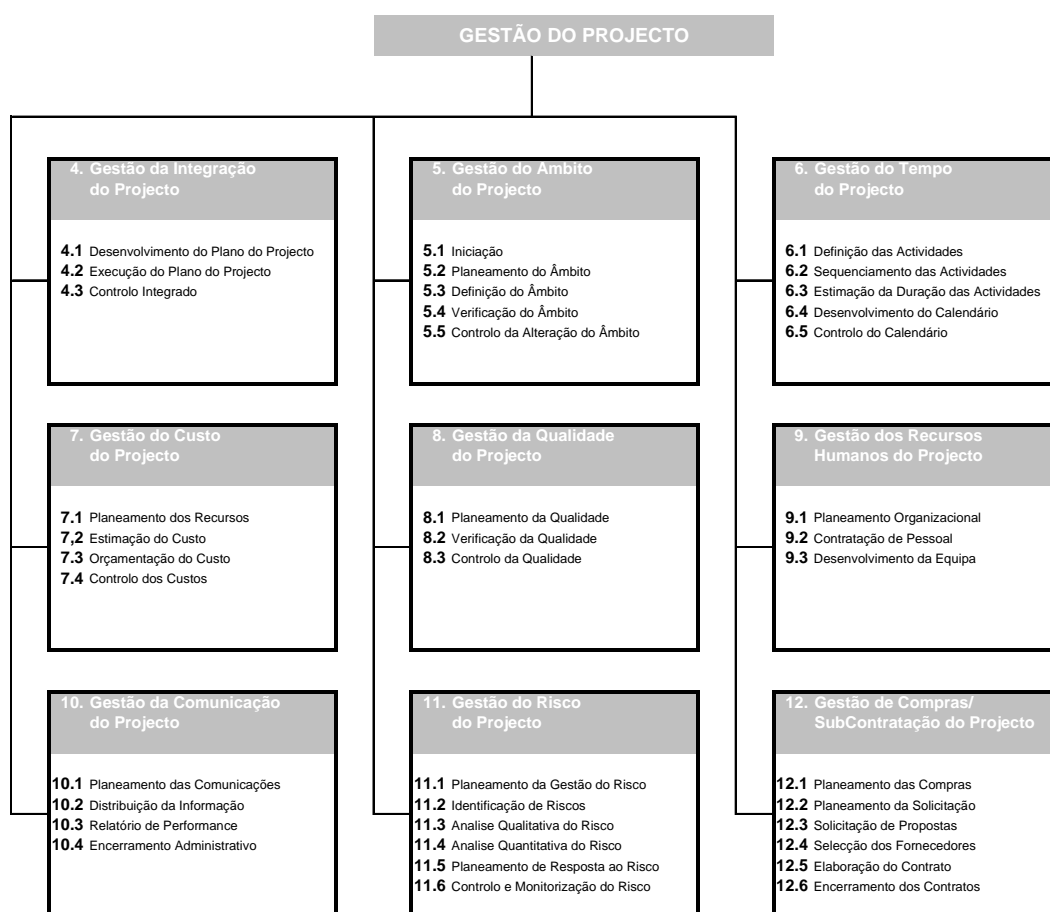


Figura 8 - Áreas do Conhecimento [PMI '00]

A Gestão da Integração descreve os processos que lidam com a gestão integrada do plano de projecto. Os processos que pertencem a esta área de conhecimento são responsáveis pelo desenvolvimento do plano integrado do projecto, pela execução das componentes do plano e pela constante monitorização do plano do projecto através da análise da sua execução.

A Gestão do Âmbito descreve os processos necessários para a definição de todas as tarefas a realizar no projecto para a conclusão com sucesso do projecto. Os processos inerentes a esta área do conhecimento tratam da iniciação do projecto, da sua definição, através da análise do plano estratégico da organização, do produto ou serviço a desenvolver e da informação histórica existente. É este processo que dá autorização para o início ou continuação para a próxima fase do projecto. Nele também é feito o planeamento do âmbito do projecto, tendo como base o produto ou serviço a desenvolver e as restrições existentes. Este planeamento visa elaborar um documento com a especificação do produto ou serviço a desenvolver e em que condições. Este documento deverá depois obter a aprovação dos *stakeholders*. O processo de definição do âmbito do projecto tem como objectivo subdividir os objectivos do documento previamente elaborado em subcomponentes de forma a facilitar a sua gestão. O processo de verificação do âmbito tem como objectivo não só a verificação do âmbito do próprio projecto, mas também a respectiva concordância de todos os *stakeholders*. Por fim, temos o processo de controlo de alterações ao âmbito do projecto que visa identificar os factores que alteram o âmbito do projecto e gerir as alterações ao âmbito de forma a que estas sejam acordadas com os *stakeholders*.

A Gestão do Tempo descreve, como o próprio nome indica, os processos necessários à gestão do tempo, nomeadamente através da definição das actividades necessárias para o alcance dos objectivos. Uma vez definidas as actividades é necessário identificar o respectivo sequenciamento e interdependências entre actividades. Existe também o processo que efectua o calculo de estimativas de duração de cada uma das actividades a desenvolver. Após a conclusão de todos estes processos é possível então desenvolver um calendário preciso para a execução das tarefas. O processo de controlo desta área do conhecimento visa monitorizar alterações ao calendário previamente estabelecido.

A Gestão do Custo descreve os processos necessários à conclusão do projecto respeitando o orçamento estipulado através do planeamento dos recursos identificados e respectiva orçamentação. Os processos de planeamento desta área de conhecimento visam elaborar um documento de necessidades da quantidade e do tipo de recursos necessários para a execução das actividades e calcular uma estimativa do custo a despendar com os recursos necessários para a execução da actividade. Adicionalmente, existe um outro processo que visa alocar os custos calculados às actividades a desenvolver. O processo de controlo de custos verifica a ocorrência de alterações nas actividades e nos recursos que têm influência sobre o valor orçamentado.

A Gestão da Qualidade descreve os processos necessários para assegurar que o projecto atinja os níveis de qualidade acordados, através do planeamento, garantia e controlo da qualidade do projecto. O planeamento da qualidade visa identificar quais os *standards* de qualidade relevantes para o projecto e de que forma é que estes serão atingidos. A verificação da qualidade é uma tarefa de avaliação constante do projecto e das suas actividades que garante a satisfação dos requisitos de qualidade preestabelecidos. O controlo da qualidade monitoriza os resultados das actividades e sugere correcções com o objectivo de atingir os *standards* de qualidade planeados.

A Gestão de Recursos Humanos descreve os processos necessários à gestão eficiente dos recursos humanos no projecto, através do planeamento de recursos necessários, da contratação e do desenvolvimento da equipa de trabalho. O planeamento organizacional identifica, documenta e atribui tarefas e responsabilidades do projecto aos elementos da equipa. O processo de aquisição de pessoal visa seleccionar as pessoas com o perfil correcto para a execução das várias tarefas do projecto. O desenvolvimento da equipa de trabalho visa melhorar as qualificações dos elementos da equipa através da formação contínua e, também, da melhoria do trabalho em equipa.

A Gestão da Comunicação descreve os processos necessários à eficiente divulgação do projecto e da sua evolução. Tal é atingido através de um planeamento da comunicação e relatórios periódicos a divulgar a todos os *stakeholders*. O planeamento da comunicação é feito através do levantamento das necessidades de informação de todos os intervenientes do projecto, da periodicidade com que a informação é necessária e do meio em como esta será divulgada. O processo de distribuição de informação visa organizar a distribuição de informação por todos os

stakeholders do projecto de forma ordenada e periódica. Adicionalmente, é também necessário desenvolver e distribuir os relatórios de desempenho e execução do projecto, com o objectivo de manter todos os *stakeholders* a par do desenvolvimento do projecto. No final, é necessário documentar e disseminar a informação necessária para se efectuar a conclusão do projecto no seu global ou de uma das suas fases.

A Gestão de Risco descreve os processos responsáveis pela identificação, análise e resposta aos riscos e oportunidades inerentes ao projecto. Inicialmente deve ser feito um planeamento sobre como gerir os riscos e as oportunidades associadas ao projecto. É então gerado um plano sobre como devem ser abordadas e planeadas as tarefas de gestão de risco em cada processo. Deve também ser feita a identificação dos riscos e oportunidades que poderão afectar o projecto, assim como as suas características, e uma análise qualitativa e quantitativa dos risco e oportunidades identificados. A análise qualitativa identifica quais os tipos de riscos e de oportunidades que poderão afectar o projecto e, como tal, influencia os objectivos do projecto, atribuindo-lhes prioridades diferentes. A análise quantitativa deverá identificar também a probabilidade de um determinado risco acontecer, assim como a sua periodicidade e respectivo impacte no projecto. Deve então ser desenvolvido um plano de acções a aplicar cada tipo de risco ou de oportunidade onde estarão descritas as técnicas e os procedimentos a adoptar para minimizar o impacte do risco sobre o projecto, ou para capitalizar a ocorrência de uma oportunidade e respectiva influência sobre o projecto. A monitorização e controlo constante do risco visa não só aplicar o plano de resposta ao risco mas também identificar novos riscos e oportunidades e avaliar a eficácia das medidas tomadas para o tratamento do risco ou aproveitamento das oportunidades.

A Gestão de Compras e Subcontratação descreve os processos necessários para a aquisição de bens e serviços externos à organização executante⁴. O planeamento das compras ou subcontratação determina quais os recursos ou tarefas que serão fornecidas ou executadas por organizações externas e quando é que estas serão efectuadas. Deve também ser feito um plano mais específico, por bem ou tarefa, identificando quais os seus requisitos e os potenciais fornecedores. Uma vez desenvolvido este plano é necessário obter propostas de diferentes fornecedores para a satisfação do plano. Estas propostas serão mais tarde avaliadas,

seleccionando-se o fornecedor daquele bem ou tarefa. A gestão do relacionamento existente com o fornecedor é de vital importância para o projecto uma vez que passa a existir uma certa dependência deste. A finalizar deve ser concluído o acordo com o fornecedor e resolvidas todas as questões pendentes.

Dado o grau de incerteza que os projectos, que por serem únicos, encerram em si, as organizações têm optado por dividi-los em diferentes fases. Desta forma é mais fácil a sua gestão e controlo, uma vez que à conclusão de cada fase estão associadas a conclusão de uma ou mais tarefas e entrega dos respectivos *deliverables*⁵. No final de cada fase são analisados os resultados das tarefas concluídas assim como o seu custo, quer monetário quer temporal. Tal análise vai permitir tomar a decisão de continuação para a fase seguinte ou de execução de eventuais acções de correcção preventivas de forma a não exceder os prazos e orçamentos estabelecidos.

A união de todas as fases de um projecto chama-se ciclo de vida do projecto. Existem diferentes ciclos de vida de projecto mediante a natureza do projecto, seja este a construção de um edifício, o desenvolvimento de um novo medicamento ou o desenvolvimento de uma aplicação de *software*. Associado a cada um destes modelos de ciclos de vida do projecto existe, normalmente, um número diferente de fases constituídas por diferentes tarefas.

Cada fase do ciclo de vida do projecto segue o esquema da Figura 7, ou seja, em cada fase, são executados os processos referentes à iniciação da fase. A repetição dos processos de iniciação no início de cada fase ajudam a manter o projecto orientado às necessidades para o qual este foi concebido para resolver.

De notar que os grupos de processos estão interligados entre si pelas entradas e saídas de informação. As entradas são documentos, estejam estes em formato físico ou digital, sobre os quais o processo vai trabalhar, e as saídas são os documentos resultantes do trabalho efectuado no processo. Aos mecanismos utilizados sobre as entradas para serem produzidas as saídas costuma-se designar por técnicas e ferramentas.

⁴ *performing organization*

Após a conclusão do processo de iniciação, são executados os processos de planeamento. Como se pode verificar pela Figura 7, o planeamento é uma tarefa permanente, pois o seu resultado serve de entrada para os processos de execução que, por sua vez, são alvo de um controlo permanente. Caso existam algumas falhas no processo de execução o resultado deste controlo pode dar origem a um novo planeamento, como por exemplo, o atraso na conclusão de uma fase. Tal acontecimento pode fazer com que seja necessário um novo planeamento relativo à necessidade de recursos materiais e humanos adicionais para o cumprimento do calendário inicial, ou a renegociação do calendário com vista a um ajustamento temporal ou orçamental. Tal interacção entre os processos de planeamento, de execução e de controlo levam a que estes sejam modelados como tarefas iterativas.

A finalizar existem os processos de encerramento cujo objectivo é garantir que os objectivos foram atingidos. Estes visam também reunir e avaliar informação sobre a conclusão da fase ou sobre o contrato de fornecimento de serviços ou bens em questão de forma a gerar e disseminar os respectivos relatórios. São também gerados relatórios sobre o conhecimento adquirido na fase ou no relacionamento com o fornecedor de serviços ou bens para consulta futura noutra fase, contrato ou projecto.

Uma vez que cada projecto é composto por um conjunto de processos e que cada um deles está classificado como pertencente a um determinado grupo de processos e a uma determinada área do conhecimento, é então possível estabelecer o mapeamento entre grupos de processos e áreas do conhecimento. Este mapeamento é demonstrado no esquema da Tabela 5.

⁵ *deliverable* – item ou resultado de alguma forma verificável, tangível ou mensurável que tem de ser produzido para completar um projecto ou parte de um projecto.

Grupos de Processos Áreas de Conhecimento	Iniciação	Planeamento	Execução	Controlo	Encerramento
4. Gestão da Integração do Projecto		4.1 Desenvolvimento do Plano do Projecto	4.2 Execução do Plano do Projecto	4.3 Controlo Integrado	
5. Gestão do Âmbito do Projecto	5.1 Iniciação	5.2 Planeamento do Âmbito 5.3 Definição do Âmbito		5.4 Verificação do Âmbito 5.5 Controlo da Alteração do Âmbito	
6. Gestão do Tempo do Projecto		6.1 Definição das Actividades 6.2 Sequenciamento das Actividades 6.3 Estimativa da Duração das Actividades 6.4 Desenvolvimento do Calendário		6.5 Controlo do Calendário	
7. Gestão do Custo do Projecto		7.1 Planeamento dos Recursos 7.2 Estimativa do Custo 7.3 Orçamentação do Custo		7.4 Controlo dos Custos	
8. Gestão da Qualidade do Projecto		8.1 Planeamento da Qualidade	8.2 Verificação da Qualidade	8.3 Controlo da Qualidade	
9. Gestão dos Recursos Humanos do Projecto		9.1 Planeamento Organizacional 9.2 Contratação de Pessoal	9.3 Desenvolvimento da Equipa		
10. Gestão da Comunicação do Projecto		10.1 Planeamento das Comunicações	10.2 Distribuição da Informação	10.3 Relatório de Performance	10.4 Encerramento Administrativo
11. Gestão do Risco do Projecto		11.1 Planeamento da Gestão do Risco 11.2 Identificação de Riscos 11.3 Análise Qualitativa do Risco 11.4 Análise Quantitativa do Risco 11.5 Planeamento de Resposta ao Risco		11.6 Controlo e Monitorização do Risco	
12. Gestão de Compras/ SubContratação do Projecto		12.1 Planeamento das Compras 12.2 Planeamento da Solicitação	12.3 Solicitação de Propostas 12.4 Seleção dos Fornecedores 12.5 Elaboração do Contrato		12.6 Encerramento dos Contratos

Tabela 5 - Mapeamento entre Áreas do Conhecimento e Grupos de Processos [PMI '00]

5.1.3 *Framework* Aplicada à Gestão de Projectos – A Visão Europeia

A *Association for Project Management* (APM), membro britânico da *International Project Management Association* (IPMA), desenvolveu nos últimos anos o seu *APM Body of Knowledge* (BoK) [APM '00a, '00b, Willis '95], no qual são abordados os diferentes tópicos que um gestor de projecto deve conhecer e dominar para realizar uma gestão de projectos eficaz.

O BoK está subdividido em sete tópicos essenciais: tópicos gerais, *framework* estratégica, controlo, características técnicas, comerciais, estrutura organizacional e pessoas.

Nos tópicos gerais são esclarecidos termos como Gestão de Projectos, Gestão de Programas (*portfolio* de projectos interdependentes) e Contexto do Projecto.

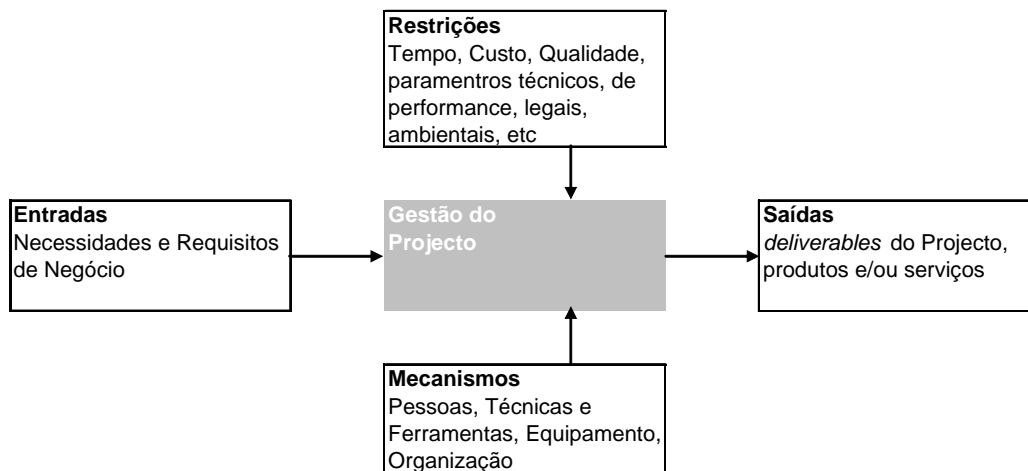


Figura 9 - Processo de Gestão de Projectos [APM '00a]

A *framework* estratégica aborda um conjunto de tópicos que ajudarão à gestão eficiente e efectiva do projecto. Aqui devem ser definidos os critérios de sucesso do projecto e desenvolvido o plano de gestão do projecto. A Gestão de Valor, o Risco e a Qualidade são também alguns dos tópicos a abordar na *framework* estratégica. A finalizar as preocupações com Saúde, Segurança e Ambiente devem também ser contempladas para a correcta gestão do projecto.

Na secção de Controlo são planeadas, avaliadas, monitorizadas e corrigidas todas as acções desempenhadas no desenvolvimento do Projecto.

A gestão da definição técnica do projecto tem grande influência sobre o sucesso ou insucesso do projecto, uma vez que devemos garantir que:

- os requisitos do projecto estão bem definidos;
- a base técnica está em conformidade com as exigências requeridas a nível de qualidade, segurança e ambiente e outros *standards*;
- existe uma estratégia tecnológica;

- o desenho está otimizado;
- as estimativas de trabalho necessário estão correctas;
- a implementação está a ser gerida;
- existe fase de testes;
- a entrega é bem gerida.

As características comerciais associadas a cada projecto são de extrema importância para o sucesso do projecto. Tópicos como justificação do projecto, gestão do financiamento do projecto, gestão imagem do projecto, gestão de subcontratados e assuntos legais são aqui abordados. O desenho do ciclo de vida do projecto, o seu desenvolvimento, implementação, entrega, avaliação pós-entrega, alocação de recursos, e definição de equipas, perfis e responsabilidades são outros dos tópicos importantes a abordar na estrutura organizacional.

A gestão das pessoas, que costuma ser uma das tarefas mais difíceis, é também de vital importância, identificando-se alguns pontos que deverão ser abordados no desenvolvimento do plano do projecto :

- comunicação com os *stakeholders*;
- trabalho de equipa;
- liderança;
- gestão de conflitos;
- negociação;
- gestão de pessoal.

Fazendo uma análise comparativa entre o PMBoK e o APM BoK, o PMBoK é uma *framework* aplicada à gestão de projectos que se tornou um *standard de facto*, pelo simples motivo desta definir uma metodologia específica para a gestão de projectos na qual são identificadas quais as tarefas a desempenhar em cada fase e os métodos e técnicas associados à gestão de cada uma dessas tarefas.

O APM BoK apresenta apenas os tópicos chave a ter em atenção na gestão de um projecto, não identificando necessariamente um princípio, meio e fim para o projecto. Levanta apenas preocupações e faz algumas sugestões sobre o que deve ser feito em cada um dos tópicos, sugerindo, até, a leitura de artigos científicos complementares sobre o tópico em questão.

5.2 Gestão de Projecto de *Data Warehousing*

O desenvolvimento de um projecto para um sistema de *data warehousing* é em tudo semelhante ao desenvolvimento de qualquer outro projecto, o que leva a que, caso se pretenda maximizar a probabilidade de sucesso e minimizar os riscos inerentes ao projecto, se deva adoptar uma boa gestão do projecto.

Uma das principais razões para a tão falada taxa de insucesso no desenvolvimento de projectos de *data warehousing* prende-se com o facto de o desenvolvimento de tais sistemas não ter sido alvo de um planeamento cuidado. A ausência ou o mau planeamento dos projectos de *data warehousing* leva a que estes projectos sejam usualmente de curta duração. Tais projectos estão normalmente atrasados, suborçamentados e os resultados são de fraca qualidade, não correspondendo assim às expectativas dos utilizadores [Adelman et al. '00].

No entanto, existem algumas particularidades num projecto de *data warehousing* que podem levantar problemas, caso sejam mal abordadas, até ao gestor de projectos mais experiente. Uma das tarefas exclusivas de um projecto de *data warehousing* é a definição dos objectivos do seu *data warehouse*. Sem uma definição de quais os objectivos a atingir com o desenvolvimento de um *data warehouse*, não há forma de avaliar o grau de sucesso do projecto nem de controlar o âmbito do projecto.

Uma vez que um *data warehouse* é a base natural de um sistema de apoio à decisão, que permite aos utilizadores efectuarem as suas interrogações sobre um repositório de dados integrados, a definição de quais as bibliotecas de interrogação a usar é uma tarefa específica e necessária ao desenvolvimento de um *data warehouse*. Uma outra tarefa exclusiva dos projectos de *data*

warehousing é a recolha, gestão e disseminação dos metadados do *data warehouse* por toda a organização, dado que sem a existência destes, dificilmente os utilizadores tirarão total partido do sistema de *data warehousing*.

A tarefa de alimentação e refrescamento de um *data warehouse* deve também ser planeada com algum cuidado, uma vez que esta tarefa tem normalmente de ser executada com uma determinada periodicidade e com restrições temporais que podem por em causa os objectivos do *data warehouse*. Se a organização é composta por filiais em diferentes pontos geográficos do mundo, e com uma janela de oportunidade limitada, o *data warehouse* poderá estar permanentemente a receber dados dos diversos sistemas fonte. Tal facto, por si só, limitará determinado tipo de consultas, como por exemplo, facturação do dia anterior. A definição de dia anterior pode ser diferente de filial para filial.

Dado que um *data warehouse* é um repositório de dados integrados, provenientes possivelmente de diferentes sistemas fontes, é necessário especificar correctamente todas as ferramentas ETI de dados dos sistemas fontes para o *data warehouse*, identificando quais e em que condições serão acedidos os sistemas fontes, os dados que serão extraídos e todos os processos de transformação que irão decorrer durante a execução destas ferramentas.

De salientar que a monitorização constante do *data warehouse* é uma tarefa de vital importância. É necessário verificar se cada alimentação do *data warehouse* terminou com sucesso e caso contrário executar as medidas correctivas planeadas para concluir a alimentação do *data warehouse*, nem que para tal seja necessário um reproprocessamento total. Deve ser analisado o ritmo de crescimento da informação armazenada no *data warehouse*, assim como, a utilização dos recursos de forma a evitar perda de desempenho que provoca insatisfação nos utilizadores.

Dadas as particularidades dos projectos de *data warehousing*, as técnicas associadas à gestão de projectos têm de ser mais restritivas, com maior ênfase no planeamento e controlo, devendo o gestor do projecto estar mais envolvido, podendo, mediante a dimensão do projecto, exigir dedicação permanente.

O planeamento de todas as fases pelas quais um projecto de *data warehousing* passa é uma tarefa essencial. A decomposição do projecto em fases e de cada fase em actividades ou tarefas torna a gestão do projecto mais simples e eficaz. No entanto, a subdivisão exagerada das tarefas em subtarefas pode contribuir com um acréscimo de trabalho administrativo e como tal atrasar a execução do projecto. Como conselho, Adelman sugere que uma tarefa que exija mais de 40 horas de trabalho deva ser subdividida. Adicionalmente, se numa tarefa existe mais que um responsável, então esta deve ser subdividida de forma a que apenas uma pessoa esteja responsável [Adelman et al. '98].

Após a identificação de quais os requisitos a satisfazer pelo *data warehouse* torna-se então possível começar a definir um calendário de tarefas com as respectivas estimativas referentes aos recursos a utilizar, sejam estes humanos ou temporais, assim como o custo estimado para a conclusão dessas tarefas. No entanto, nem sempre o desenvolvimento do calendário é feito após o levantamento dos requisitos, é comum a gerência da organização definir uma data como limite para a operacionalidade do projecto. Restrições temporais ou orçamentais tendem a limitar o grau de sucesso de um sistema de *data warehousing* [Adelman et al. '00, Kimball et al. '98].

Uma vez que o desenvolvimento de um sistema de *data warehousing* é feito de forma incremental, cada incremento pode ser considerado como um projecto independente. No entanto, deverá sempre haver um calendário global para todos esses incrementos, assim como um planeamento inicial que defina a arquitectura geral do sistema de *data warehousing* [Inmon '02, Kimball et al. '98].

Outro factor chave no desenvolvimento de um projecto de *data warehousing* são os recursos humanos. É ainda difícil encontrar dentro das organizações pessoas qualificadas nesta área. Tal facto, leva a que seja normalmente necessário recorrer a pessoal externo, o que eleva, em muito, o custo com pessoal qualificado. A formação do pessoal interno é por vezes a solução. No entanto, esta formação extraordinária requer o seu tempo, o que poderá não ser aceitável caso o limite para a conclusão do projecto não o permita. É preferível a contratação de pessoal especializado comparativamente à contratação ou alocação de pessoal que requer formação, porque um projecto de *data warehousing* é normalmente um projecto que consome bastante tempo. Como tal, a

experiência de uma pessoa especializada em determinada área (análise, modelação, ferramentas ETI) diminuirá o tempo consumido em formação, adaptação, pesquisa e na execução das tarefas. De salientar também que, um elemento altamente especializado terá já participado em diversos projectos de *data warehousing*, o que lhe terá permitido assimilar conhecimento que mais tarde será útil e como tal, cometer menos erros nas tarefas a executar. A diferença principal entre a execução de um projecto genérico e um projecto de *data warehousing* é o facto de este último não ser estático. Nele existem mudanças constantes que devem ser geridas de forma a levar o projecto à sua conclusão.

Dada a natureza de um *data warehouse*, à medida que este é desenvolvido e apresentado aos *stakeholders* do projecto, ao patrocinador, aos membros da equipa e potenciais utilizadores, irão surgir novos pedidos ou requisitos que poderão alterar o âmbito do projecto e, como tal, comprometer todo o planeamento efectuado previamente. Embora o gestor do projecto no seu planeamento inclua uma margem de manobra para lidar com problemas não esperados, a inclusão de novos requisitos poderá eliminar essa margem e comprometer seriamente prazos e custos estimados para o projecto. Caso algum novo requisito seja adicionado ao projecto o âmbito é alterado sendo assim necessário renegociar com o seu patrocinador os objectivos, os prazos e o orçamento do projecto.

O desenvolvimento de um projecto acarreta riscos, riscos esses que devem ser lidados atempadamente e eficazmente de forma a reduzir o seu impacte no projecto. Os projectos de *data warehousing* têm um grau ainda mais elevado de incerteza e risco dado que o seu desenvolvimento é relativamente recente e, como tal, ainda existem poucos gestores de projectos que tenham lidado com este novo tipo de problemas. A inexperiência dos gestores de projectos em lidar com projectos de *data warehousing* eleva ainda mais o grau de incerteza do projecto e, assim, todos os factores de risco devem ser contemplados num plano de gestão de risco juntamente com as acções a desempenhar de forma a minimizar as suas influências menos positivas no projecto. Os principais factores de risco dos projectos de *data warehousing* são:

- Perda do patrocinador, uma vez que este é o impulsionador do projecto na organização e que atribui ao projecto recursos necessários ao seu desenvolvimento.

- A não aceitação do projecto por parte dos utilizadores também comprometerá o grau de sucesso do projecto. Tal deve ser minimizado tentando incluir os utilizadores no desenvolvimento do projecto desde as fases iniciais do projecto.
- O fraco desempenho de um sistema de *data warehousing* poderá constituir mais um factor de afastamento dos utilizadores do novo sistema, daí que deve ser elaborado um plano de longo prazo para o desenvolvimento de uma arquitectura que suporte o crescimento do sistema sem comprometer o seu próprio desempenho.

Finalmente, um dos aspectos mais negligenciados nos projectos de *data warehousing* é a gestão da comunicação do projecto. Existem dois tipos de comunicação que devem ser abordados: a comunicação interna da equipa do projecto e a comunicação com os elementos externos à equipa do projecto.

A nível interno, todos os elementos da equipa devem seguir um esquema de comunicação periódica do desenvolvimento das tarefas a que estão alocados. Este esquema contempla reuniões semanais de curta duração de toda a equipa do projecto onde são abordados os desenvolvimentos e problemas que surgiram durante a semana. Estas reuniões devem estar previamente organizadas com um conjunto de tópicos a abordar. No final de cada reunião cada membro deverá receber um pequeno relatório de todas as questões abordadas e de todas as decisões tomadas de forma a haver um perfeito entendimento do rumo do projecto.

A nível externo, devem ser conduzidas apresentações com periodicidade mensal de forma a manter os patrocinadores e potenciais utilizadores informados sobre o desenvolvimento do projecto, onde são revistos os objectivos do projecto e comunicadas as dificuldades ultrapassadas, ou por ultrapassar, assim como o seu impacte a nível de orçamento e calendarização.

Acima de tudo, o gestor do projecto deve nas suas comunicações e apresentações ser honesto e apresentar o real desenvolvimento do projecto, justificando-o. Tal honestidade poderá elevar a probabilidade de sucesso do projecto uma vez que o patrocinador saberá sempre como realmente está a evoluir o projecto diminuindo assim o risco de surpresas e de um possível abortar do projecto.

5.3 Kimball e PMI

Ralph Kimball propõe uma metodologia para o desenvolvimento de SDW denominada de Ciclo de Vida Dimensional do Negócio [Kimball et al. '98]. Esta metodologia subdivide o desenvolvimento do SDW em várias tarefas, algumas desenvolvidas de forma sequencial, outras em paralelo, e outras de forma contínua. Estas últimas requerem um acompanhamento constante ao longo do projecto do SDW.

Assim pretende-se, estudar nesta secção observar até que ponto a metodologia proposta por Kimball se integra com a *framework* proposta pela PMI para a gestão de projectos. À imagem do *Project Management Body of Knowledge* (PMBok) [PMI '00], Kimball define que após a iniciação do projecto de *data warehousing* seja efectuado o seu planeamento detalhado. Este plano deverá ser desenvolvido após análise de diversos factores, de vital importância, para o sucesso do projecto, como a avaliação da necessidade do *data warehouse*, a avaliação da capacidade de adaptação e preparação da organização para o impacte do desenvolvimento de um SDW, em que são analisadas questões como a existência de um forte patrocinador da área do negócio⁶, a efectiva necessidade por parte da organização de um *data warehouse*, a existência de uma parceria entre a área de negócio e o departamento de informática, e por fim a cultura organizacional da organização em causa. Uma vez analisados e ponderados todos estes factores, através da execução do *Litmus Test* [Kimball et al. '98], decide-se se se avança ou não com o desenvolvimento do SDW.

Caso haja autorização para a iniciação do projecto, a fase seguinte é o desenvolvimento do plano do projecto de *data warehousing*. Este desenvolvimento passa primeiro pelo planeamento das diversas tarefas a executar, processos esses que segundo a classificação do PMBoK pertencem ao grupo de processos de Planeamento e que abrangem todas as áreas do conhecimento. Kimball define que a primeira tarefa, após a autorização para o desenvolvimento do projecto, seja o

⁶ Gestor com poderes executivos dentro da organização que define quais os objectivos e prazos de implementação e que garante que o projecto tem os recursos, quer financeiros quer humanos, necessários à sua implementação.

planeamento e definição do âmbito do projecto. Tais tarefas são equivalentes aos processos 5.2 e 5.3 do PMBoK [PMI '00] que também são executadas na fase inicial do projecto.

Uma vez definido o âmbito do projecto, em conjunto com o patrocinador do negócio, é necessário justificar quais os custos e investimentos necessários para o desenvolvimento do SDW, assim como, os benefícios e retornos financeiros que se obterão com a utilização do SDW. O planeamento dos custos envolvidos no projecto é uma tarefa equivalente ao processo 7.2 e 7.3 da PMBOK, que têm o nome, respectivamente, de Estimação e Orçamentação do Custo, não havendo processo equivalente para a projecção dos benefícios e retornos financeiros do projecto. De notar que, para uma efectiva projecção dos custos e investimentos envolvidos no projecto, será necessário efectuar um planeamento prévio dos recursos necessários, sejam estes materiais, financeiros ou humanos, assim como a definição das actividades a desempenhar e respectivo calendário. Todo este planeamento é abordado pelo PMBoK nos processos 6.1 – Definição das Actividades, 6.2 – Sequenciamento das Actividades, 6.3 – Estimação da Duração das Actividades e 7.1 – Planeamento dos Recursos .

A finalizar o planeamento do projecto do SDW, é desenvolvido o plano do projecto onde são discriminadas as tarefas a desenvolver, o seu escalonamento, os recursos necessários para cada tarefa, assim como, o esforço para a sua execução, a data prevista de início, data prevista de conclusão, data de conclusão efectiva e estado. O desenvolvimento do plano do projecto é abordado no processo 4.1 do PMBoK.

Após o desenvolvimento e a aprovação do plano do projecto do SDW, é dado o início do seu desenvolvimento. Existe um conjunto de tarefas que tratam do desenvolvimento do SDW, e uma tarefa paralela a todas estas, denominada de Gestão do Projecto, que acompanha o projecto do seu início até à sua conclusão. Esta tarefa é responsável pelo controlo e monitorização das tarefas que executam o plano do projecto do SDW.

A primeira tarefa associada à gestão do projecto é a reunião de arranque com todos os membros da equipa de desenvolvimento do SDW, para que sejam apresentados os objectivos e o âmbito do projecto, e discutidas as responsabilidades e tarefas de todos os membros da equipa. É também

apresentado o plano do projecto a alto nível, sem dados referentes ao esforço planeado e datas de conclusão. A finalizar o gestor do projecto deverá distribuir por todos os participantes um resumo da reunião, e toda a documentação que permita melhorar a comunicação e o entendimento entre a equipa de desenvolvimento, como por exemplo, uma lista de contactos, uma lista de responsabilidades, o plano e o âmbito do projecto, assim como, todos os documentos necessários ao registo de eventos ocorridos durante o desenvolvimento de uma tarefa.

Durante o decorrer do desenvolvimento do SDW devem ser agendadas reuniões periódicas entre a equipa de desenvolvimento, nas quais são revistas as tarefas executadas, os resultados obtidos, e distribuídas as próximas tarefas a executar. Devem também ser revistos os problemas e tarefas pendentes desde a última reunião, assim como feita a avaliação ao seu novo estado, e analisados os novos problemas e tarefas em atraso de forma a arranjar soluções. São abordadas nestas reuniões questões relacionadas com alterações ao plano ou ao âmbito e avaliado o respectivo impacto no plano original. O resultado destas reuniões é resumido num relatório a distribuir por todos os participantes. Adicionalmente, é elaborado um relatório a distribuir pelos *stakeholders* com o estado do projecto, contendo uma descrição do trabalho até então desenvolvido, as questões ainda pendentes ou eventuais alterações ao plano inicial. Este conjunto de reuniões fazem parte do processo 10.3 – Relatório de Performance, onde é analisado o estado das tarefas em desenvolvimento.

No desenvolvimento de um projecto, a definição, controlo e monitorização do âmbito do projecto são de grande importância. No entanto, quando se trata de um SDW, a diferença entre uma boa ou má gestão do âmbito do projecto tem como consequência o sucesso ou insucesso do projecto. Kimball aconselha a que o gestor do projecto tome uma de três posições quando existe um pedido de alteração do âmbito do projecto:

- Negar a inclusão de novos requisitos.
- Ajustar o âmbito do projecto de forma a incluir os novos requisitos mas eliminando outros, de forma a que o orçamento e calendário se mantenha.
- Negociar com os patrocinadores do projecto uma extensão de prazo e orçamento para a inclusão dos novos requisitos no âmbito do projecto.

Estas tarefas estão associadas aos processos 5.4 e 5.5 que são, respectivamente, de verificação do âmbito e gestão de alterações ao âmbito do projecto.

A finalizar os temas abordados na gestão de projectos, Kimball aconselha o desenvolvimento de uma estratégia de comunicação com os *stakeholders* de forma a que todos estejam informados do objectivo, do trabalho desenvolvido e das dificuldades existentes no projecto. Esta comunicação vai permitir gerir melhor as expectativas de todos e principalmente dos utilizadores finais. Esta estratégia de comunicação é abordada no processo 10.2 – Distribuição de Informação.

O desenvolvimento do SDW começa pela análise dos requisitos de negócio, onde são entrevistados os *stakeholders*, patrocinadores de negócio e utilizadores finais, com o objectivo de reunir todos os requisitos que identificarão e determinarão prioridades de desenvolvimento do SDW. A análise dos requisitos originará um documento que deverá ser aprovado pelos patrocinadores do projecto e distribuído pelos *stakeholders* identificando as áreas prioritárias no desenvolvimento do SDW.

Uma vez escolhido o assunto a incluir no SDW, o desenvolvimento do SDW reparte-se em três linhas de acção que poderão ser executadas em paralelo, a saber: a linha de acção tecnológica, a linha de acção dos dados e a linha de acção das aplicações. A linha de acção tecnológica visa analisar e desenvolver toda a arquitectura técnica que irá sustentar o desenvolvimento e funcionamento do SDW. A linha de acção dos dados tem como principais tarefas a análise dos dados a incorporar no *data warehouse*, através da análise das fontes de dados, o desenvolvimento do modelo dimensional que armazenará os dados previamente tratados e integrados, a construção do modelo físico do *data warehouse* e o desenvolvimento das ferramentas ETL que extrairão, tratarão e integrarão os dados no *data warehouse*. A linha de acção das aplicações contém as tarefas de especificação e desenvolvimento das aplicações que interagirão com o *data warehouse* e que serão usadas pelos utilizadores finais com vista à obtenção de dados provenientes deste.

A cada uma das linhas de acção propostas por Kimball pode ser aplicada a metodologia de gestão de projectos do PMBoK, como se cada linha de acção fosse um projecto.

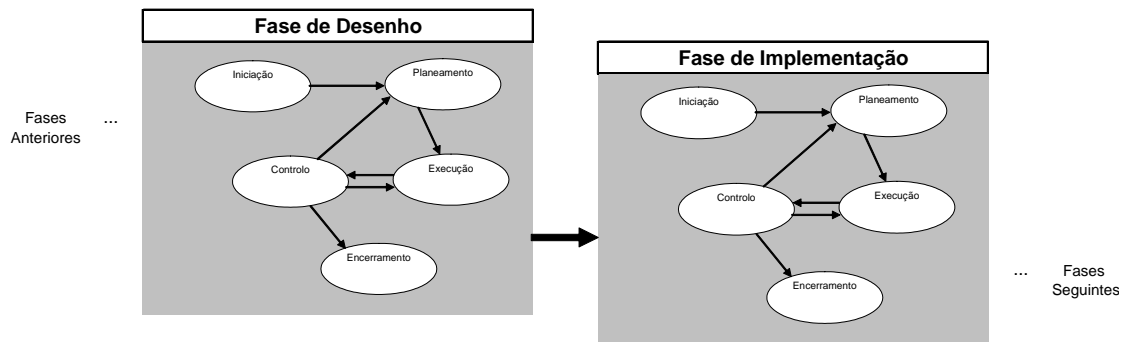


Figura 10 - Interação entre Fases [Duncan '95]

É no desenvolvimento destas três linhas de acção que são executados principalmente os processos da área de conhecimento do Risco, Qualidade e Compras e Subcontratação.

A finalizar o desenvolvimento destas três linhas de acção, existe a fase de Instalação e Arranque, onde são integrados os resultados das três linhas de acção, constituindo-se assim o SDW. Durante o desenvolvimento das três linhas de acção e na fase de Instalação e Arranque são especialmente importantes os processos de integração do projecto abordados na área de conhecimento da Gestão da Integração.

Capítulo 6

Da Teoria à Prática

6.1 Casos de Estudo

Uma vez concluída a investigação e avaliadas as propostas dos mais variados autores na área de *Data Warehousing* e de Gestão de Projectos, foi necessário validar toda a recolha de informação tendo em conta as actuais práticas no mercado nacional nesta área e avaliar quais as metodologias e técnicas usadas em projectos de *data warehousing* contrapondo com a teoria proposta por alguns dos autores estudados.

Foram seleccionadas algumas empresas que respeitassem o requisito de terem participado ou desenvolvido um projecto de *data warehousing*, ou que ainda estivessem no decorrer desse tipo de projectos. Foi estabelecido o contacto com os responsáveis das empresas seleccionadas ou dos departamentos responsáveis pelo desenvolvimento do *data warehouse*, e efectuadas entrevistas de duração prolongada aos líderes dos projectos com o intuito de melhor compreender as metodologias utilizadas no desenvolvimento de projectos de *data warehousing*. Estas entrevistas

tinham também o intuito de identificar pontos em comum no planeamento, gestão e desenvolvimento de projectos de *data warehousing*, registando ideias chave assim como regras de ouro do ponto de vista de cada líder em projectos de elevado risco como são os de *data warehousing*.

As empresas que aceitaram participar e colaborar nestas entrevistas foram a Enabler – *Solutions for Retailing* através do Gestor do Produto de *Data Warehousing* da MCH – Eng.º Paulo Faria e através da Gestora do Produto de *Data Warehousing* (RDW) da DeSpar – Dr.ª Emília Ribeiro, a Sonae Indústria através do responsável pelo departamento dos SI – Eng.º José Agostinho e a PT Inovação através do Responsável pela área de *Business Intelligence* (BI) – Eng.º João Pias.

Dada a dimensão do país e da realidade nacional da área de *data warehousing*, acreditamos que as empresas entrevistadas são uma amostra credível, ficando apenas, das áreas mais relevantes, o negócio da banca por analisar. Nesta análise estão representados o retalho, através das entrevistas realizadas aos gestores dos diversos produtos da Enabler – *Solutions for Retailing*, a indústria, com a entrevista ao responsável dos SI da Sonae Indústria, e as telecomunicações, através da entrevista com o responsável pela área de BI da PT Inovação.

Empresa	Projecto	Quadros	Funções
Enabler, <i>Solutions for Retailing</i>	DW da MCH	Eng.º Paulo Faria	Gestor de Produto
	DW da DeSpar	Dr.ª Emília Ribeiro	Gestor do Produto
Sonae Indústria	DW interno	Eng.º José Agostinho	Responsável pelo Departamento de SI
PT Inovação	Vários projectos	Eng.º João Pias	Responsável pela Área de BI

Tabela 6 - Empresas entrevistadas

As entrevistas realizadas seguiram as fases da metodologia proposta por Kimball [Kimball et al. '98], que foi dada a conhecer aos entrevistados e onde foram efectuadas e respondidas questões relacionadas com os projectos de *data warehousing* em que os entrevistados estiveram envolvidos. Quer a Enabler quer a PT Inovação são empresas que prestam serviços na área da informática, e

como tal envolvidas em diversos projectos de *data warehousing* em simultâneo. A Sonae Indústria desenvolveu o seu *data warehouse* recorrendo à especialização e conhecimento do departamento de SI da empresa não tendo sido necessário escolher um parceiro de desenvolvimento externo à empresa.

Dada a heterogeneidade das empresas entrevistadas, torna-se necessário assentar um ponto prévio de importância significativa na análise dos projectos em que estas estão ou estiveram envolvidas. A Sonae Indústria desenvolveu o seu *data warehouse* internamente, recorrendo ao departamento de SI, sendo possível analisar e comentar todos os passos envolvidos num projecto desta natureza. A Enabler e a PT Inovação, como empresas de prestação de serviços na área de *Data Warehousing*, estão apenas sujeitas a análise dos métodos utilizados na gestão e desenvolvimento do *data warehouse*, não sendo possível analisar a parte inicial do projecto relativa à definição do projecto, sua justificação, constituição da equipa interna, e identificação do patrocinador. Estas tarefas são efectuadas, previamente à atribuição do projecto à empresa de desenvolvimento, por parte do cliente.

A abordagem seguida para a apresentação dos resultados das entrevistas deriva da apresentação da metodologia mais consensual e mais utilizada na área de *Data Warehousing* – o Ciclo de Vida Dimensional [Kimball et al. '98] – em conjunto com as observações efectuadas por cada entrevistado quanto à necessidade, relevância e ponto de vista crítico das fases e tarefas que esta metodologia contém.

Recorrendo ao diagrama previamente estudado da metodologia (Figura 6), será feito um estudo a cada uma das fases, às tarefas que estas envolvem e respectiva análise por parte dos entrevistados, começando pelo Planeamento do Projecto, Análise de Requisitos, Gestão do Projecto, desenvolvimento do *data warehouse* através das 3 linhas de acção, Instalação e Arranque, e Manutenção e Evolução.

6.2 Planeamento do Projecto

6.2.1 Definição do Projecto

De acordo com Ralph Kimball, os projectos de *data warehousing* têm como início um pedido ou uma necessidade de desenvolvimento que será satisfeita com a existência de um *data warehouse*. Sendo assim, torna-se necessário definir o projecto, identificando a origem desse pedido e necessidade. Por norma, o facto que despoleta o pedido ou necessidade de um *data warehouse* está classificado segundo três vertentes:

- Pedido por parte de um departamento de negócio – Quando um departamento tem necessidade de melhor informação e como tal requisita o desenvolvimento de um *data warehouse* para a satisfação dos seus requisitos
- Pedido por parte de vários departamentos do negócio – Quando existem vários departamentos a requisitar o desenvolvimento de um *data warehouse*.
- Em busca da necessidade ou pedido – Acontece quando é decidida a construção de um *data warehouse* sem ter havido necessariamente um pedido por parte de algum departamento da organização.

Das entrevistas efectuadas, apenas a Sonae Indústria passou por esta fase, uma vez que a Enabler e a PT Inovação, como empresas prestadores de serviços, não estão presentes nesta fase inicial do projecto.

De acordo com José Agostinho, o desenvolvimento do *data warehouse* surgiu com a necessidade de obtenção de uma Consolidação das Vendas e Análise da Performance das Linhas de Encomenda de sistemas fontes dispersos geograficamente (Europa, América, África). Este pedido proveio da administração da empresa que sentiu necessidade de obtenção de resultados globais com maior celeridade e exactidão.

Uma vez definido o projecto e identificada a origem do pedido, é necessário avaliar a prontidão da empresa, identificar a motivação do negócio e analisar a cultura organizacional. Estas tarefas foram executadas no decorrer do planeamento do projecto de *data warehousing* da Sonae Indústria concluindo-se que a empresa estava preparada para iniciar o projecto, que a principal motivação para tal projecto era a necessidade, por parte da administração, de informação consolidada referente às vendas de todas as suas filiais como forma de melhor avaliar a performance destas, e detectar atempadamente variações e tendências nas vendas dos seus produtos. O facto da Sonae Indústria ser composta por empresas deslocadas geograficamente potenciou a necessidade de informação consolidada uma vez que cada filial detinha fisicamente e geria os seus SI. No entanto, o ERP que suportava os SI era comum a todas as filiais, tendo sido desenvolvido pelo departamento de SI da Sonae Indústria, o que facilitaria a extracção e integração dos dados. De salientar, que a partir de 2000, a estratégia referente aos SI da Sonae Indústria alterou-se, tendo sido planeado um sistema centralizado baseado no SAP⁷ para substituir o SI distribuído existente desenvolvido internamente. Neste momento, ainda a meio do processo de transição, existem filiais com o ERP antigo, com ambos, ou apenas com o sistema SAP.

Concluída a análise aos factores atrás enunciados e havendo condições para se iniciar o projecto deve ser feita uma primeira definição do âmbito do projecto. Esta definição deve ser uma tarefa executada pelo departamento de SI em parceria com representantes da área do negócio. Em paralelo, deve também ser construída uma justificação do projecto onde são identificadas quais as suas necessidades e quais os benefícios que se pretendem alcançar.

No caso da Sonae Indústria, o âmbito do projecto estava previamente definido pela necessidade identificada e o custo do projecto foi incluído no orçamento anual do departamento de SI.

Os projectos de *data warehousing* estudados na Enabler são diferentes, o *data warehouse* da MCH é considerado como um projecto de desenvolvimento iterativo e constante, onde são definidos no início de cada ano civil pelo cliente, com a participação da equipa da Enabler, os objectivos a alcançar nesse ano, nomeadamente novas áreas a integrar no *data warehouse*, novas funcionalidades, entre outros requisitos. O *data warehouse* da DeSpar está inserido no *package* da

⁷ <http://www.sap.com>

Retek⁸, Retek *Data Warehouse* (RDW), e é considerado como um projecto de manutenção evolutiva, em que o cliente vai solicitando novas funcionalidades que vão sendo implementadas ou reajustadas tendo em conta o conceito de *data warehouse* da Retek.

A PT Inovação tem em mãos diversos projectos de *data warehousing*. A maioria destes projectos foram adjudicados à PT Inovação através de concurso público no qual está incluído um caderno de encargos que define qual o âmbito geral do projecto e quais os requisitos a satisfazer.

6.2.2 Calendarização do Projecto

A calendarização dos projectos é determinada tendo em conta dois factores, as tarefas a executar e o pessoal disponível para as executar. Se houver limitação de tempo imposta pelo patrocinador ou cliente, os recursos despendidos para a conclusão do projecto vão ter de ser superiores. Caso não exista limite de tempo definido, este é determinado pelos recursos humanos existentes e o tempo necessário para a conclusão das tarefas com esses recursos humanos.

Nos projectos estudados existem ambas as versões, no caso da Sonae Indústria em que o desenvolvimento é interno, executado pelo Departamento de SI, qualquer novo requisito ou iteração do *data warehouse* é calendarizado tendo em conta os recursos existentes, ou seja, a equipa de *data warehousing*. Anualmente é desenvolvido um plano de actividades e uma lista de objectivos, nos quais são descritos os projectos a desenvolver nesse ano. N entanto, durante o decorrer do ano pode haver alterações às estratégias definidas no plano de actividades.

No projecto de *data warehousing* da MCH, o cliente, em parceria com a Enabler, também define anualmente os objectivos e as áreas a atingir assim como o orçamento necessário para a execução dessas tarefas. Esta particularidade apenas é possível dado o excelente relacionamento entre a Enabler e a MCH.

⁸ <http://www.retek.com>

Dada a particularidade do projecto de *data warehousing* da DeSpar, já enunciada anteriormente, a calendarização do projecto é diferente dos restantes projectos estudados. Este projecto não tem um fim planeado à vista, uma vez que uma manutenção evolutiva, como o próprio nome indica, não tem fim. No entanto, cada novo requisito ou pedido do cliente pode ser tratado como um projecto, ou mini-projecto, uma vez que a sua duração pode oscilar entre poucos dias a 6 meses. Em projectos de longa duração, existe uma maior formalidade, com troca de documentação de parte a parte onde é documentado o pedido, o orçamento para o desenvolvimento do pedido, a avaliação ao orçamento, entre outros documentos. Em mini-projectos, os pedidos são normalmente enviados via *email* e satisfeitos, e só posteriormente é que é tratada a documentação inerente ao registo do pedido e do orçamento. Foi adoptado este tipo de procedimento tendo em conta a base de confiança existente entre as empresas e com o objectivo de minimizar o tempo de disponibilização do novo requisito.

Quanto à PT Inovação, esta baseia-se no calendário especificado no caderno de encargos do projecto que lhes foi adjudicado. O caderno de encargos define, entre outras coisas, os objectivos e prazos do projecto. Existe também um calendário “macro” que é depois especificado pelo comité técnico e dividido em fases. À conclusão de cada fase está associado um conjunto de *deliverables*.

Independentemente da abordagem de cada uma das empresas para a calendarização dos seus projectos, o tempo estimado para a conclusão de uma determinada tarefa é baseado no conhecimento adquirido de projectos anteriores ou de tarefas semelhantes.

6.2.3 Orçamentação do Projecto

A finalizar o desenvolvimento do plano do projecto é necessário identificar quais as necessidades financeiras do projecto.

Ao Departamento de SI da Sonae Indústria é atribuído um orçamento anual para a execução das suas actividades. À medida que os projectos de *data warehousing* vão surgindo, sejam de dimensão reduzida ou de complexidade baixa, são executados e o orçamento anual absorve esse

custo. Se o projecto for complexo ou moroso, é elaborado um orçamento que é apresentado à filial que efectuou o pedido para suportar o seu desenvolvimento.

Como já foi referido anteriormente, o desenvolvimento do projecto de *data warehousing* da MCH está sujeito a um plano anual onde são definidos os objectivos a atingir assim como o orçamento necessário para atingir esses objectivos.

No caso do projecto de *data warehousing* da DeSpar, em cada projecto é desenvolvido um orçamento para a execução do projecto, onde são discriminadas as tarefas e o custo que estas tarefas acarretam. Este orçamento tem de ser aprovado pelo cliente para que o projecto tenha início.

A PT Inovação participa em vários projectos de desenvolvimento de *data warehouse*, adjudicados maioritariamente através de concursos públicos, onde estão previamente definidos o calendário e o orçamento para execução do projecto. No entanto, no decorrer do projecto este orçamento é ajustado conforme as necessidades do projecto pelo comité de “alto nível”.

6.2.4 Constituição da Equipa de *Data Warehousing*

Obtida a aprovação do projecto por parte da administração da organização e concedidos os recursos requeridos para a conclusão do projecto, torna-se necessário constituir as equipas de trabalho mediante as áreas de desenvolvimento e com especial atenção às aptidões necessárias para a conclusão com sucesso das tarefas existentes.

Kimball define um conjunto de perfis necessários numa equipa de desenvolvimento de um *data warehouse*, no entanto, mediante a dimensão de cada projecto, poderão haver várias pessoas com o mesmo perfil a desempenhar uma tarefa ou até um pessoa que desempenhe funções de perfis variados. Os perfis identificados por Kimball são:

- Analista do Negócio.

- Arquitecto de Sistemas de Dados.
- Administrador das Bases de Dados do *data warehouse*.
- Projectista do sistema da área de retenção.
- Programadores das aplicações dos utilizadores finais.
- Formador em *data warehousing*.

A dimensão das equipas de *data warehousing* das empresas analisadas varia entre 2 a 15 elementos. Vários factores contribuem para a dimensão da equipa. Se o desenvolvimento do *data warehouse* é externo, irão necessariamente existir pelo menos duas equipas, a equipa do cliente e a equipa da empresa que faz o desenvolvimento.

No caso das empresas prestadoras de serviços, se estas apenas fizerem o desenvolvimento do *data warehouse* (processos ETI, Desenho Lógico e Físico) a equipa será de reduzidas dimensões. Se a equipa de desenvolvimento do *data warehouse* também tiver a seu cargo o desenvolvimento das aplicações que vão interagir com o *data warehouse* então a dimensão da equipa será superior. Caso a empresa esteja envolvida em variados projectos de *data warehousing*, a equipa da empresa prestadora de serviços poderá atingir uma dimensão considerável para o plano nacional, cerca de 15 pessoas. De notar, que estas pessoas poderão não estar todas envolvidas ao mesmo tempo no mesmo projecto, uma vez que poderão estar a decorrer vários projectos ao mesmo tempo e cada um deles ter uma equipa reduzida.

Como exemplo, a Enabler tem uma equipa de *data warehousing* para cada um dos projectos existentes, neste caso o *data warehouse* da MCH e o *data warehouse* da DeSpar, liderados respectivamente por Paulo Faria e Emilia Ribeiro. No *data warehouse* da MCH, existem 3 equipas com tarefas diferentes, a equipa de suporte, a equipa de desenvolvimento do *data warehouse* e a equipa de desenvolvimento de explorações.

A equipa de suporte tem a seu cargo o suporte aplicacional, constituído por 3 pessoas, que têm como principal tarefa a monitorização dos processos de alimentação do *data warehouse*, através da utilização de ferramentas que relatam o estado dos processos. Esta equipa era mantida por elementos da Enabler até ao ano de 2003, e desde então passou a ser composta por pessoal do

cliente. O suporte técnico é assegurado por um elemento da equipa de desenvolvimento da Enabler que tem como principal tarefa diagnosticar e efectuar pequenas correcções em processos problemáticos.

A equipa de desenvolvimento é composta por cinco elementos residentes, e mediante a complexidade e a quantidade de projectos existentes pode ser alargada recrutando elementos temporários ou realocando elementos da equipa de desenvolvimento de explorações para a equipa de desenvolvimento do *data warehouse* uma vez que estas pessoas têm bons conhecimentos na área de *Data Warehousing*. Actualmente, a equipa é composta por cerca de 10 a 12 pessoas a trabalhar no desenvolvimento de projectos de *data warehousing*. Sendo a de desenvolvimento do *data warehouse* composta por Analistas Programadores e Gestores de Projecto e a de desenvolvimento de explorações por 5 a 6 analistas programadores com bons conhecimentos de *data warehousing* que desenvolvem aplicações de consulta ao *data warehouse*.

O outro projecto da Enabler que também foi analisado refere-se ao *data warehouse* da DeSpar. É um projecto de dimensão inferior ao da MCH e com um perfil ligeiramente diferente. A diferença entre os dois projectos nota-se também na dimensão da equipa do projecto. Esta equipa é composta apenas por 2 elementos residentes, podendo, mediante a sobrecarga de trabalho, ser complementada temporariamente por mais um elemento externo. Numa equipa desta dimensão os elementos da equipa efectuem as tarefas de diferentes perfis, como por exemplo, análise, desenho e desenvolvimento. No projecto de *data warehousing* da DeSpar existe também uma equipa do cliente que é responsável pelo desenvolvimento das aplicações de exploração e *reporting*. Esta equipa integra também um administrador de bases de dados que gere o *data warehouse*.

Na PT Inovação, a equipa de desenvolvimento do *data warehouse* é composta por cerca de meia dúzia de elementos residentes, estando neste momento com uma dimensão total de 15 elementos dada a diversidade de projectos em que estão a participar. Estes elementos são contratados temporariamente mediante a exigência dos projectos em curso.

No entanto, os projectos em que a PT Inovação participou ou ainda tem em curso costumam estar organizados da seguinte forma:

- Comité de “alto nível” – composto por representantes ou elementos da administração da PT Inovação e do Cliente.
- Comité técnico – composto pelo líder de projecto de ambas as empresas mais um número reduzido de elementos, normalmente entre 4 a 5 pessoas. Este comité é responsável pela gestão corrente do projecto, calendarização do projecto e comunicação com a administração.
- Equipa de desenvolvimento do *data warehouse* – equipa que desenvolve os procedimentos de extracção e *staging* do *data warehouse*.
- Equipa de desenvolvimento de aplicações e explorações – equipa que desenvolve as aplicações de consulta, exploração e *reporting* do *data warehouse*.
- Mediante a dimensão do projecto é aconselhada a existência de um equipa autónoma à equipa de desenvolvimento, para fazer o acompanhamento, auditoria e controle de custos do projecto e que reporte os resultados ao comité de “alto nível” criado para o projecto.

Por último, analisando o projecto de *data warehousing* da Sonae Indústria, conclui-se que a equipa de desenvolvimento é composta por 4 Analistas Programadores e a equipa da filial local é composta no mínimo por 2 pessoas, responsável local e responsável local pelas infra-estruturas tecnológicas. O responsável local pelas infra-estruturas tecnológicas, quando incorporado num projecto de *data warehousing*, tem como tarefa o suporte aplicacional e a transmissão de problemas à equipa central que também efectua o suporte deste elemento.

Em qualquer um dos casos analisados o patrocinador do projecto é, por norma, o cliente. Mesmo no caso do desenvolvimento interno da Sonae Indústria, à excepção da iteração inicial em que o patrocinador foi a administração, o patrocinador é o cliente, vulgarmente denominado por filial uma vez que a Sonae Indústria é composta por várias filiais.

A definição do líder do projecto, no caso da Enabler e da PT Inovação, é efectuada em conjunto com a empresa cliente, havendo um líder da parte da empresa que faz o desenvolvimento e um líder da equipa da empresa cliente.

Uma vez alocadas as pessoas às equipas e definidas as tarefas a executar, é necessário desenvolver um plano detalhado e integrado do projecto de *data warehousing*. Este plano deve identificar as tarefas a executar no desenvolvimento do projecto, em conjunto com os recursos necessários à execução da tarefa, identificação da data de início, data prevista de conclusão, data efectiva de conclusão, estado da tarefa, dependências e estado da tarefa. Kimball aconselha à utilização dos métodos e sistemas de *software*, já existentes nas empresas, para elaboração do plano do projecto, lembrando que o tempo gasto na elaboração de um plano de projecto apenas fará sentido se este for mantido e actualizado no decorrer do projecto.

6.3 Gestão do Projecto

A gestão de um projecto é essencial para a sua conclusão com sucesso. Kimball sugere um conjunto de técnicas para a gestão eficiente de um projecto que uma vez adoptadas aumentam a probabilidade de sucesso diminuindo o risco inerente a um projecto desta natureza. Uma vez desenvolvido o plano do projecto é necessário implementá-lo e geri-lo. A gestão do projecto deve começar por uma reunião inicial com a equipa de *data warehousing*, onde são explanados os objectivos a atingir com o desenvolvimento do projecto, assim como o seu estado corrente e quais os passos a executar segundo um calendário pré-definido e quais as tarefas e responsabilidades atribuídas a cada elemento da equipa.

6.3.1 Monitorização do Projecto

Um dos principais factores que contribuem para uma boa gestão de projecto é a documentação. A inexistência de documentação torna o projecto ingovernável, daí que, todas as tarefas devem estar documentadas, todas as decisões devem estar registadas, assim como todos problemas e

respectivas soluções. Só assim é possível saber o que há para fazer, o que foi feito, porque é que foi feito daquela forma e quais os problemas existentes e ainda por resolver.

Existem duas técnicas conhecidas e utilizadas para monitorizar o estado de um projecto, através de reuniões periódicas com os elementos da equipa de *data warehousing* e através de relatórios periódicos, com o ponto da situação do projecto, que devem ser dados a conhecer não só à equipa mas também aos patrocinadores.

Nas reuniões periódicas deverão ser revistas as tarefas em curso, especialmente as que estão atrasadas relativamente ao prazo esperado de conclusão, devem ser debatidos os problemas que surgirão desde a última reunião. O resultado destas reuniões deverá ser reflectido no plano do projecto elaborado previamente, onde são actualizados os estados de cada tarefa, ou com a data de conclusão ou assinalando o seu atraso.

Os relatórios periódicos deverão reflectir o progresso do projecto, resumindo as reuniões periódicas, ou seja, assinalando tarefas concluídas, plano de trabalhos para o período seguinte e identificando as tarefas atrasadas. Estes relatórios deverão ser breves, garantindo assim que estes sejam realmente lidos por aqueles a quem se destinam.

As metodologias seguidas pelas empresas analisadas para a gestão de projectos não diferem muito do sugerido, nem diferem muito entre elas. A Sonae Indústria informou que dada a proximidade e dimensão da equipa de *data warehousing* prevalece a comunicação informal entre os membros da equipa, no entanto, existem reuniões periódicas entre a equipa onde é feito o ponto de situação do projecto e com a filial cliente e responsável local pelas infra-estruturas tecnológicas. Estas reuniões podem ser locais, por videoconferência ou usando o *netmeeting*.

No caso do projecto de *data warehousing* da MCH, a Enabler efectua reuniões de arranque do projecto, quer reuniões internas com toda a equipa de *data warehousing*, quer reuniões externas com o cliente e utilizadores finais. Caso se justifique, existem reuniões periódicas entre toda a equipa de *data warehousing*. Novamente, a proximidade dos elementos da equipa facilita o intercâmbio de ideias e sugestões na resolução de problemas. Uma vez que o *data warehouse* da

MCH é desenvolvido pela Enabler, que é uma empresa externa à MCH, é necessário haver comunicação entre a equipa da MCH e a equipa da Enabler. Esta comunicação deveria ser estabelecida apenas entre os gestores do projecto de cada equipa, mas, dada a relação de proximidade entre as duas equipas, tal não acontece, levando a que a comunicação entre elementos das duas equipas se proceda à excepção da norma.

No outro projecto de *data warehousing* da Enabler, o *data warehouse* da DeSpar, as alterações a nível de metodologia utilizada para a gestão do projecto não são muitas, a comunicação entre elementos da equipa é informal, existem reuniões periódicas entre a equipa, caso se justifique. De lembrar que esta equipa é constituída por 2 a 3 pessoas, o ponto de situação com o cliente é efectuado quinzenalmente.

Na PT Inovação, a metodologia adoptada para a gestão do projecto deriva das normas impostas pela certificação da qualidade, no entanto, os pontos em comum com as empresas anteriormente abordadas são, comunicação informal entre elementos da equipa de *data warehousing*, comunicação entre a equipa da PT da PT Inovação e a equipa do Cliente, via gestores do projecto de *data warehousing*. Também existe troca de documentação formal ou informal, nomeadamente via *email*, entre elementos da equipa interna e com o cliente. São efectuadas reuniões periódicas semanais para avaliar as tarefas pendentes, sendo documentadas todas as decisões tomadas e respectivas justificações para uma melhor gestão do projecto.

6.3.2 Gestão do Âmbito

Um dos maiores factores de risco inerentes a um projecto de *data warehousing* é a gestão do seu âmbito. Uma das garantias que podemos ter relativamente ao âmbito de um projecto de *data warehousing* é que este vai se alterar, tal afirmação é preconizada por Kimball. Partindo do princípio que o âmbito do projecto pode alterar-se, é necessário identificar métodos de gestão que nos permitam gerir essas mudanças e concluir com sucesso o projecto. As alterações ao âmbito do projecto estão inerentes à descoberta de novos requisitos durante o processo de desenvolvimento, ou devido a novas funcionalidades exigidas pelos utilizadores finais ou patrocinadores.

Kimball advoga uma de três posições perante uma alteração no âmbito do projecto, nomeadamente:

- A não alteração do âmbito.
- Ajustamento do âmbito com o pressuposto de que este não alterará o esforço necessário à conclusão do projecto.
- Alargamento do âmbito com a respectiva alteração de orçamento e calendário.

Os gestores de projectos de *data warehousing* foram questionados sobre quais os procedimentos adoptados no caso de uma alteração do âmbito do projecto e as respostas quanto à importância da gestão do âmbito e respectivas tarefas de gestão variam um pouco.

Na Sonae Indústria há duas abordagens diferentes, quando a alteração do âmbito é mínima este é alargado de forma a incluir essa alteração ou novo pedido. Quando a alteração ao âmbito do projecto é significativa, é sugerida uma 2ª fase do projecto a começar após a conclusão da 1ª fase e tendo em linha de conta os outros projectos previamente calendarizados. No *data warehouse* da MCH, segundo Paulo Faria, a determinação do âmbito de um projecto é “regra de ouro” num projecto de *data warehousing*. Na negociação e aprovação do âmbito do projecto são determinados todos os requisitos a satisfazer, especialmente o que fazer quanto à qualidade dos dados fonte, limpar ou não. Se houver dificuldade na definição do âmbito do projecto, ou caso as tarefas a executar sejam de elevada complexidade, a Enabler promove *workshops* e entrevistas com os utilizadores finais, com o cliente, e com os intervenientes do projecto para analisar com o maior detalhe possível os requisitos do cliente. No entanto, se houver lugar à alteração do âmbito do projecto no decorrer deste, esta alteração terá de ter o acordo do cliente e consequente renegociação do calendário e orçamento inicial.

6.3.3 Plano de Comunicação

Um dos pontos críticos na execução de um projecto de *data warehousing* é a comunicação. Existem três grupos com requisitos diferentes de informação que necessitam de ser informados sobre o decorrer do projecto, são eles, a equipa de *data warehousing*, os patrocinadores e os utilizadores finais. Por isso, o gestor do projecto deve estabelecer um plano de comunicação com estes três grupos de forma a conseguir não só gerir o projecto como concluí-lo com sucesso.

Como já foi enunciado anteriormente, a melhor forma de comunicar com a equipa de *data warehousing* é através de reuniões periódicas onde são discutidas e avaliadas as tarefas até então executadas, apresentadas as tarefas pendentes, suas razões e equacionadas possíveis soluções, à medida que se vai fazendo o acompanhamento do plano de projecto previamente desenvolvido. No entanto, se a equipa de *data warehousing* for numerosa, poderá recorrer-se a outras formas de comunicação, nomeadamente através da distribuição de relatórios de progresso do projecto ou das actas das reuniões periódicas com os responsáveis de cada grupo dentro da equipa de *data warehousing*.

Adicionalmente, devem ser também agendadas reuniões periódicas, com um intervalo de tempo superior ao das reuniões da equipa de *data warehousing*, envolvendo os patrocinadores do projecto, uma vez que estes são parte interessada na conclusão com sucesso do projecto e co-responsáveis pelo projecto perante a administração da organização. Kimball sugere como preferíveis as reuniões cara-a-cara ao simples envio de relatórios de progresso do projecto. Acima de tudo, os patrocinadores do projecto deverão ter uma visão precisa do estado do projecto como forma não só de co-responsabilização mas também como forma de integração no projecto.

Por último, mas também de grande relevância para o sucesso do projecto, deve ser desenvolvido um plano de comunicações que vise manter os utilizadores finais informados sobre o novo sistema que está a ser desenvolvido e que será utilizado por estes. A gestão das expectativas dos utilizadores finais é de extrema importância para o projecto. Se os utilizadores finais sentirem que fazem parte do projecto, que as suas sugestões são ouvidas pela equipa de desenvolvimento e que o projecto corresponde às suas expectativas, maior será aceitação e utilização deste novo sistema,

contribuindo significativamente para o seu sucesso. Por outro lado, se os utilizadores finais são colocados à parte do desenvolvimento do projecto, se o seu conhecimento do negócio não é aproveitado no desenvolvimento do novo sistema ou se este não corresponde às suas expectativas, o novo sistema terá menor aceitação e provavelmente menos utilização. Tal facto seria mais um factor a contribuir para o insucesso do projecto.

O estudo da aplicação do plano de comunicações sugerido por Kimball sofre do mesmo problema anteriormente descrito quanto ao tipo de empresas estudadas e respectivos projectos. Como estas são maioritariamente empresas de prestação de serviços, que são contratadas por empresas que querem ver desenvolvido um *data warehouse*, a comunicação com os utilizadores é, salvo raras excepções, da responsabilidade da equipa do cliente e, como tal, de difícil análise. No entanto, uma das excepções, é o projecto de *data warehousing* da MCH, onde a equipa de desenvolvimento teve a necessidade de organizar *workshops* com os utilizadores finais como forma não só de os motivar e fazer com que se sintam parte do processo, mas também para que estes contribuam na análise do projecto. A PT Inovação por seu lado participa nas sessões de esclarecimento promovidas pelo cliente aos seus utilizadores finais como forma de perceber melhor as suas necessidades e também de promover a motivação e cooperação nos utilizadores finais.

A comunicação entre a equipa de *data warehousing* e com os patrocinadores de negócio foram já descritas no tópico referente à monitorização do projecto. De sublinhar que é extremamente importante que o patrocinador de negócio tome conhecimento do real andamento do projecto.

6.3.4 Qualidade

Um dos temas importantes para qualquer projecto é a qualidade do produto e do trabalho desenvolvido. Este tema não é abordado pela metodologia de Kimball, mas foi abordado nas entrevistas realizadas como forma de melhor entender como é feito o controlo de qualidade do trabalho desenvolvido e como é que o cliente tem garantias do produto que lhe está a ser entregue.

Um dos pontos favoráveis às empresas que fornecem serviços é a obtenção da certificação da qualidade. Esta certificação garante que existam normas, procedimentos, documentação e qualidade no desenvolvimento dos projectos em que a empresa certificada está envolvida. Esta certificação serve também como forma de diferenciação aquando da subcontratação do desenvolvimento de um projecto de *data warehousing*.

Das empresas estudadas, apenas a PT Inovação tinha a certificação de qualidade, daí que em todos os projectos em que a PT Inovação participa há um pressuposto de garantia incorporada no projecto. A metodologia de gestão de projectos é a imposta pelo sistema da qualidade da Empresa.

A Enabler, embora não sendo uma empresa que tenha a certificação de qualidade, segue um guia interno de desenvolvimento de *software*, horizontal a toda a empresa, que determina normas para a nomenclatura dos objectos, onde os guardar e como os guardar. Relativamente a projectos de *data warehousing*, está em desenvolvimento um guia interno com um conjunto de normas e procedimentos à adoptar no desenvolvimento deste tipo de projectos. Foi já feita uma análise sobre o impacte da aplicação desse guia interno ao *data warehouse* da MCH e concluiu-se que as alterações a implementar eram incomportáveis, daí que, apenas os novos projectos em que a Enabler participa tendem a incorporar essas normas para facilitar o crescimento do *data warehouse* e organização interna da equipa de *data warehousing*.

A qualidade do desenvolvimento no projecto de *data warehousing* da MCH é garantida pelo seguimento dos princípios de desenvolvimento da Enabler e por auditorias ao código desenvolvido. Essas auditorias são efectuadas por amostra e por pessoal externo à equipa de *data warehousing*. De notar que alterações ao código do que está em produção implica a notificação do Gestor do Produto. A certificação do projecto é assegurada pelo cumprimento das normas internas para a entrega de um Projecto, entre elas a verificação de uma *checklist* de pontos que o Projecto tem de cumprir.

No projecto de *data warehousing* da DeSpar, além das normas de qualidade já enunciadas no outro projecto em que a Enabler participa, são também seguidas as normas de desenvolvimento aconselhadas pela *Retek*.

6.3.5 Factores Críticos

Uma das boas práticas seguidas na gestão de projectos consiste em identificar os factores críticos para o projecto e definir as acções preventivas de forma a minimizar a sua importância, e as acções correctivas que reduzirão o impacto do problema no evoluir do projecto. No entanto, cada projecto tem o seu conjunto de factores críticos. Poderão existir alguns pontos em comum, mas a sua importância difere de projecto para projecto. Uma das questões colocadas aos entrevistados foi referente aos factores críticos que na sua opinião maior influência têm sobre o desenrolar do projecto.

José Agostinho, responsável pelo Departamento de SI da Sonae Indústria, identifica como um risco para o projecto o possível incumprimento das tarefas ou prazos por parte dos responsáveis dos sistemas de informação locais, elegendo três factores críticos como os mais importantes no projecto de *data warehousing* que têm sido desenvolvidos:

- Comunicação , dada a diversidade de nacionalidades das filiais da Sonae Indústria.
- Planeamento do Projecto, dada a sua influência no orçamento e calendarização de actividades do Departamento de SI.
- Qualidade dos Dados dos Sistemas Fonte, dada a existência de um processo de migração dos Sistemas Transaccionais.

Paulo Faria, gestor de produto *data warehousing* da MCH, por seu lado identifica dois factores críticos que necessitam de especial atenção no desenvolvimento do *data warehouse*:

- Qualidade dos Dados dos Sistema Fonte.
- Envolvimento dos Utilizadores Finais.

Emilia Ribeiro, gestora de produto de *data warehousing* da DeSpar, identifica a alimentação dos novos *Data Mart* com dados históricos como o factor crítico do projecto em que está inserida.

Por seu lado, João Pias, responsável pela área BI da PT Inovação, identifica como factores críticos a Especificação e Análise de Requisitos, a Comunicação com o Cliente e o bom ambiente entre a equipa de *data warehousing* interna, ou seja, o factor humano.

Resumindo, os pontos em comum referidos pelos entrevistados como factores críticos dos projecto de *data warehousing* em que estiveram, ou em que estão inseridos, são a comunicação com o cliente e a qualidade dos dados dos sistemas fonte.

6.3.6 Dificuldades

Para ser possível uma melhor compreensão da complexidade de um projecto de *data warehousing*, questionou-se os entrevistados sobre quais as maiores dificuldades ou as dificuldade mais comuns que tiveram de enfrentar num projecto.

Na Sonae Indústria, as dificuldades mais comuns são:

- A Gestão do Âmbito dos vários projectos de *data warehousing*.
- As inúmeras fontes e sua heterogeneidade.
- O acompanhamento tecnológico (*Upgrades*, Mudanças de versões).
- As dificuldade em manterem-se actualizados dada a escassez do tempo.

Estas dificuldades advêm, principalmente, do facto do desenvolvimento do *data warehouse* estar a cargo do Departamento de SI que é interno à Sonae Indústria. Torna-se difícil à equipa de *data warehousing* justificar ao cliente, que é uma filial ou departamento da Sonae Indústria, a não inclusão de novos requisitos durante o decorrer do projecto. Uma vez que a Sonae Indústria é composta por várias filiais e que está a decorrer um processo de migração dos sistemas

transaccionais, a heterogeneidade e variedade de fontes é uma dificuldade acrescida ao projecto de *data warehousing*. A subdimensão da equipa de *data warehousing* aliado à limitação orçamental imposta ao Departamento de SI, dificulta a actualização e acompanhamento dos elementos da equipa aos constantes desenvolvimentos na área de *Data Warehousing*.

Na Enabler, no caso do projecto de *data warehousing* da MCH, as maiores dificuldades prendem-se com o reprocessamento de tarefas uma vez que é necessário ter em atenção as respectivas dependências. Um exemplo da necessidade de efectuar um reprocessamento seria a falta de alguns dados de uma loja que obriga ao reprocessamento do dia todo. Como consequência do reprocessamento de tarefas, existe também o respectivo impacte na comunicação com os utilizadores finais, e na performance do Sistema dada a sobrecarga imposta pelo reprocessamento.

No projecto de *data warehousing* da DeSpar as dificuldades prendem-se com:

- A comunicação com o cliente dada a diversidade das línguas. O cliente é italiano e a equipa de *data warehousing* do cliente é composta por pessoas de diferentes nacionalidades.
- O facto da máquina de desenvolvimento estar localizada no cliente implica a utilização de uma linha de comunicações que é falível.
- O horário de trabalho implica que seja necessário efectuar alguns procedimentos e testes na hora de intervalo para almoço e a partir das 20h até as 6h Italianas.
- O cliente às vezes vê o *data warehouse* como um sistema transaccional, exigindo algumas funcionalidades típicas de um sistema transaccional.

Na PT Inovação, dada a diversidade de projectos em que participam, as dificuldades mais comuns nos projectos de *data warehousing* são:

- O relacionamento com o Cliente.
- A Análise e Especificação de Requisitos.
- O negócio muito dinâmico, ou seja, em constante mudança.
- O acesso às fontes de dados, uma vez que estas são “propriedade” de outras equipas.

Estas dificuldades estão maioritariamente relacionadas com o facto do cliente ser de nacionalidade diferente, dificultando não só o relacionamento com o cliente, como as tarefas de análise e especificação de requisitos. A constante mudança do negócio implica também uma constante actualização por parte dos analistas.

Como se pode verificar, as dificuldades variam de projecto para projecto sendo que a dificuldade mais comum continua ser o relacionamento com o cliente.

6.4 Recolha e Análise de Requisitos

6.4.1 Entrevistas e *Workshops*

Kimball sugere o uso de um conjunto de técnicas para a recolha de requisitos que um *data warehouse* deverá suportar, em particular o uso de entrevistas com os patrocinadores, com os gestores e com os utilizadores finais para melhor compreender as suas necessidades e as suas expectativas quanto ao projecto em si. Quando o número de possíveis participantes é grande as entrevistas individuais consomem muito tempo tornando-se impraticáveis. Sendo assim, é possível através de sessões de análise e esclarecimento, *workshops*, recolher e analisar requisitos de um conjunto diversificado de pessoas, e ao mesmo tempo cativá-los e incentivá-los à participação no projecto.

6.4.2 Sucesso ou Insucesso

Após a recolha e análise de requisitos, Kimball sugere que se determinem algumas métricas para avaliar o sucesso ou insucesso de um projecto. O conjunto de métricas sugeridas são diversas e variam desde métricas quantitativas, métricas de utilização, métricas sobre o nível de serviço e métricas associadas ao impacto no negócio.

Uma vez questionados, os entrevistados, na sua grande maioria, determina o sucesso ou insucesso do projecto através da monitorização da utilização ou não do *data warehouse*. A Enabler envia aos seus clientes questionários que pretendem avaliar o projecto quanto ao grau de satisfação e ao grau de sucesso do *data warehouse*. Estes questionários são respondidos pelos gestores do projecto de *data warehousing* do cliente e pelos seus utilizadores finais. De salientar que a administração da Enabler definiu como objectivo uma média de 3,6 nas avaliações dos seus clientes para todos os projectos concluídos.

No caso do *data warehouse* da DeSpar, a média das avaliações efectuadas pelo cliente determina uma classificação de 3,6 para os projectos até então concluídos.

6.5 Implementação

Uma vez desenvolvido o plano do projecto de *data warehousing*, constituídas as equipas e atribuídas as tarefas, a construção do SDW deve ser executado em paralelo, no qual o desenvolvimento da arquitectura que suporta o *data warehouse*, o desenho lógico e físico do *data warehouse* e o desenho e implementação das aplicações para os utilizadores finais constituem as três linhas de acção.

6.5.1 Linha de Acção Tecnológica

A linha de acção tecnológica visa o desenvolvimento da arquitectura técnica necessária ao suporte de todas as actividades do *data warehouse*. Nesta linha é elaborado o desenho da arquitectura técnica e posteriormente, implementada a arquitectura planeada através da selecção do *hardware* e *software* necessário ao desenvolvimento do *data warehouse*.

O *data warehouse* da Sonae Indústria, dado o modesto número de registos que são adicionados diariamente, tem ainda uma dimensão relativamente pequena. Mas a adição recente de um novo *Data Mart* na área de *Stocks* implicará o crescimento do *data warehouse* em cerca de 200%. No

entanto, este crescimento, não põe em causa a actual arquitectura do *data warehouse* uma vez que este está já programado para albergar 5 anos de dados históricos. Cada novo requisito que surge, e é implementado no *data warehouse*, é analisado ao pormenor, projectando-se o n.º de acessos, n.º de registos criados, n.º de pessoas que vão aceder por dia de forma a estimar os requisitos técnicos necessários, e estimado o espaço em disco para uma utilização a dois anos que é automaticamente alocado. O *data warehouse* está armazenado numa base de dados Oracle 8⁹ e está a ser planeado o *upgrade* para a versão 9, uma vez que o suporte à versão 8 vai ser descontinuado pela *Oracle*. Este *upgrade* vai exigir do Departamento de SI o desenvolvimento de soluções para manter a conectividade com o Oracle 7 dado que com o Oracle 9 tal deixa de ser possível. Como mecanismo de melhoria de desempenho a equipa de *data warehousing* optou pelo particionamento da base de dados, diminuindo assim o tempo de resposta de *queries* associadas a um determinado conjunto de valores. A título exemplificativo, o particionamento por região do mundo da tabela de Vendas vai acelerar o tempo de resposta das *queries* associadas à análise de vendas de uma região específica do globo.

O projecto de *data warehousing* da MCH tem também a participação de uma equipa do cliente vocacionada para a gestão de infra-estruturas que tem como principal preocupação a gestão da arquitectura do *data warehouse*. Esta equipa é composta por Administradores de Bases de Dados e Engenheiros de Sistemas. No entanto, segundo opinião de Paulo Faria, ainda não é dada a relevância necessária à gestão do sistema e das infra-estruturas do *data warehouse*. A existência de um Administrador de Bases de Dados a tempo inteiro na equipa de *data warehousing* da Enabler é relativamente recente. O SGBD que suporta o *data warehouse* neste momento é o Oracle 9, tendo vindo a ser actualizado desde o Oracle 7 utilizado até ao ano 2000 e Oracle 8 até ao ano 2003.

No projecto da DeSpar, as infra-estruturas são geridas pelo cliente. A equipa de *data warehousing* da Enabler por diversas vezes sentiu a necessidade de ajudar o Administrador da Bases de Dados do cliente em problemas com o *data warehouse* assim como na implementação de novas funcionalidades no sistema. O SGBD em que o *data warehouse* está armazenado é o Oracle 8.

⁹ <http://www.oracle.com>

6.5.2 Linha de Acção de Dados

Nesta linha de acção devem ser desenhados e construídos os modelos de dados que suportam as necessidades de informação dos utilizadores finais. São também desenvolvidos os procedimentos de extracção, transformação, limpeza e alimentação dos dados dos sistemas fontes para o *data warehouse*.

Quando um projecto de *data warehousing* é desenvolvido internamente, todas as tarefas são da responsabilidade da equipa de *data warehousing* do projecto ou do Departamento de SI. Quando o projecto é subcontratado a empresas externas, todas as combinações são possíveis, quanto às linhas de acção da responsabilidade da empresa externa. No entanto, o que é mais comum é associar o desenvolvimento dos modelos de dados e das ferramentas de ETI à empresa externa, deixando a gestão da infra-estrutura tecnológica a cargo do cliente. O desenvolvimento das aplicações dos utilizadores finais poderão ser da responsabilidade da empresa externa, ou, caso haja um *software* específico, o cliente poderá adquiri-lo e parametrizá-lo com a ajuda dessa empresa.

O projecto de *data warehousing* da Sonae Indústria foi desenvolvido internamente, daí que as três linhas de acção são da responsabilidade do Departamento de SI. Relativamente à linha de acção dos dados, a Sonae Indústria adoptou um modelo de dados relacional para suportar o seu *data warehouse*. Os dados são extraídos de fontes de dados heterogéneas, dado o processo de transição do antigo ERP para o SAP, para a *Staging Area* onde são convertidos e agregados. No processo de extracção é efectuada uma verificação sintáctica dos dados, enquanto que a verificação semântica ocorre na fase de testes do *data warehouse*. Todo o desenvolvimento das ferramentas ETI foi feito internamente.

Questionado sobre a justificação para a ocorrência de erros, José Agostinho identifica três principais razões e as medidas possíveis para os corrigir: Erros de dados nos sistemas fonte, onde sempre que é possível estes são corrigidos. Caso o cliente não queira que esses valores sejam corrigidos nas fontes, mas sim no *data warehouse*, é obrigado a responsabilizar-se por essa alteração. No *data warehouse* a alteração é efectuada, mas sempre com a possibilidade de se

saber o historial das alterações. Erros no *data warehouse*, normalmente são provenientes de fórmulas mal implementadas. Erros nos relatórios provenientes de *bugs* de acesso aos dados no *data warehouse*. A consequência da detecção de erros por parte dos utilizadores finais é a descredibilização do *data warehouse* e consequente falta de confiança na informação por este prestada.

No projecto de *data warehousing* da MCH, a Enabler desenvolveu todas as ferramentas de extracção, transformação, limpeza e alimentação dos dados para o *data warehouse*. As ferramentas ETI deste projecto funcionam como uma caixa negra, sendo apenas necessário parametrizá-la para aceitar novos fluxos de dados ou novas fontes de dados, desde que estes sejam de um formato que já tenha interface, caso contrário é necessário desenvolver o interface para este novo formato. Estas ferramentas têm sido alvo de melhorias à medida que decorre cada novo projecto.

Em cada processo de alimentação, os dados são extraídos das fontes e integrados no ODS que depois alimentam os diversos *Data Marts* existentes. Existem posteriormente um conjunto de aplicações e sistemas que consulta a informação armazenada nos *Data Marts*. Dada a quantidade elevada de dados que o *data warehouse* recebe diariamente, estes têm de ser ordenados por prioridade e com as condicionantes da janela de oportunidade de cada sistema fonte. Desta forma, os dados mais críticos são extraídos durante a noite, e os restantes vão sendo extraídos e alimentados no *data warehouse* durante o resto do dia. De salientar que o período de indisponibilidade do *data warehouse* é bastante curto.

O gestor do produto de *data warehousing* da MCH da Enabler foi questionado sobre a utilização ou não de chaves de substituição. Neste momento o *data warehouse* da MCH já usa chaves de substituição em duas dimensões dos seus mais recentes projectos, permitindo assim ter acesso ao historial da codificação de determinados atributos, não havendo planos para alterar os *Data Marts* já existentes, mas existindo a vontade de implementar chaves de substituição nos novos projectos. A alteração de todos os *Data Marts* existentes para a utilização de chaves de substituição é uma tarefa incomportável para a Enabler uma vez que tal acção exigiria grandes recursos, quer financeiros quer temporais, que o cliente não estaria disposto a assumir.

O *data warehouse* da DeSpar, o *Retek Data Warehouse* (RDW), está integrado num *package* da Retek, o que à partida facilita bastante a tarefa de ETI, uma vez que a grande maioria das fontes de dados pertencem ao mesmo *package*. Segundo Emilia Ribeiro, os dados a incorporar no RDW provêm do Retek Merchandising System (RMS), são extraídos durante a noite e integrados nas tabelas atómicas sendo depois alimentados nos Data Marts. Em casos excepcionais, é desenvolvido o interface para extracção de dados de sistemas fonte que não são integradas no RMS. A utilização de chaves de substituição nos *Data Marts* é comum embora sejam de difícil compreensão para o utilizador final.

Nos projectos em que João Pias da PT Inovação participou, ou está a participar, existem ligeiras diferenças quanto ao acesso aos dados e ao desenvolvimento e uso das ferramentas ETI. Existem projectos em que os processos de ETI são relativamente simples, dado conhecimento interno da fonte de dados e a facilidade de acesso ao sistema. Existem também alguns projectos que são muito complexos dada a heterogeneidade das fontes e o seu difícil acesso, dado que estas estão em mudança, ou porque o processo de actualização ou a arquitectura do sistemas transaccionais estão a ser alterados.

Os interfaces de acesso às fontes são por norma desenvolvidos pelas equipas que gerem as fontes de dados, não só pelo factor de conhecerem bem a fonte, mas também por poderem identificar melhor qual o período para fazer a extracção. Desta forma a equipa de *data warehousing* não tem necessidade de conhecer a fonte, nem de aceder a ela, apenas tem de comunicar à equipa que gere a fonte quais os dados que pretende, quando os pretende e onde deseja que eles sejam colocados. A partir desse momento, caso seja necessário, é efectuada a transformação e limpeza dos dados para serem, posteriormente, integrados no *data warehouse*. As chaves de substituição são utilizadas sempre, como forma de manter um histórico de actualizações dos dados operacionais, com o objectivo de isolar a tabela de factos de actualizações nas chaves primárias do negócio, e também como forma de optimização de espaço de armazenamento no *data warehouse*.

João Pias alerta para a questão da gestão dos metadados como uma tarefa de importância vital para um *data warehouse*. Nos projectos em que esteve envolvido, João Pias advoga um sistema de

gestão de metadados técnicos e de negócio como forma de melhorar o entendimento do SDW, quer por parte da equipa de *data warehousing*, quer por parte dos utilizadores finais. Os metadados técnicos clarificam o historial de transformações efectuadas durante os processos de ETI, e os metadados de negócio esclarecem qual o significado de cada atributo para o negócio.

6.5.3 Linha de Acção das Aplicações

Associado a esta linha de acção estão as tarefas de desenho e desenvolvimento de aplicações de acesso ao *data warehouse* para consulta e emissão de relatórios. Estas aplicações serão usadas pelos utilizadores finais que tirarão partido dos dados armazenados no *data warehouse* para suportarem as suas decisões.

Na Sonae Indústria, o *data warehouse* serviu inicialmente com o fonte de informação a mais de mil relatórios que são construídos e distribuídos diariamente através do *Lotus Notes*¹⁰. Cada utilizador final tem, conforme a sua posição, acesso a um subconjunto dos mapas existentes. Existem milhares de clientes que diariamente consultam os relatórios publicados. Neste momento existe um conjunto restrito de utilizadores que utiliza a aplicação *Business Objects*¹¹ para efectuar as suas consultas Ad-hoc.

O desenvolvimento das ferramentas de exploração do *data warehouse* da MCH está a cargo de uma equipa própria, composta por 5 a 6 pessoas e que desenvolvem em *Business Objects*, *Crystal Reports* e *Holos*.

No caso do *data warehouse* da DeSpar, o departamento de informática do cliente até à cerca de um ano atrás era o único com acesso a este, e era o responsável pela construção de relatórios que eram distribuídos pelos “verdadeiros” utilizadores do *data warehouse*. Neste momento, já existe

¹⁰ <http://www.lotus.com/notes>

¹¹ <http://www.businessobjects.com>

um conjunto restrito de utilizadores que acede ao *data warehouse* através da utilização da aplicação *Microstrategy*¹².

6.6 Instalação e Entrega do Projecto

Na fase de Instalação e Entrega do Projecto é feita a convergência das três linhas de acção. Esta é uma tarefa de elevada complexidade, uma vez que grande parte do desenvolvimento é feito em paralelo. É nesta fase que o desenvolvimento da arquitectura do técnico, do *data warehouse* e das aplicações de exploração, uma vez "unidas" formam o Sistema de *Data Warehousing*. Nesta fase existe também a componente de formação e suporte que ajudará a integrar os utilizadores finais no projecto de *data warehousing* e assim contribuir para o seu sucesso.

Na Sonae Indústria a metodologia de desenvolvimento utilizada é interna, mas comum a outras abordagens, começando pela Análise, seguido do Desenho, respectivo Desenvolvimento, a fase de Testes, e a culminar a Publicação e Formação. Fazendo uma breve comparação desta metodologia com a de Kimball, poderemos dizer que estamos perante a fase de Publicação e Formação. É nesta fase que são disponibilizados o novo conjunto de relatórios que tiram partido dos desenvolvimentos recentes, e é dada formação ao responsável local pelas infra-estruturas tecnológicas no âmbito do suporte ao *data warehouse*.

Relativamente aos projectos de *data warehousing* em que a Enabler participa, os procedimentos adoptados para a entrega de um projecto já foram mencionados na análise à qualidade dos projectos, isto é, os projectos na Enabler antes de serem entregues e sujeitos à aprovação do cliente têm que seguir um rigoroso controlo onde é verificado se o projecto cumpre uma lista de requisitos. A formação dos utilizadores finais é normalmente assegurada pelo cliente, no entanto, é necessário dar formação à equipa do cliente, normalmente o departamento de informática, para que esta se sinta à vontade não só no desenvolvimento de relatórios e de ferramentas de exploração, mas também para dar formação aos seus utilizadores finais.

¹² <http://www.microstrategy.com>

6.7 Manutenção e Crescimento

A conclusão de um projecto de *data warehousing* pode não ser a conclusão do desenvolvimento de um *data warehouse*, mas apenas a conclusão de uma de muitas iterações. No entanto, é a partir da conclusão da primeira iteração que a equipa de *data warehousing* deve começar a prestar atenção à manutenção e ao crescimento do *data warehouse* uma vez que este não é estático. A manutenção inclui tarefas como monitorização do *data warehouse*, de todos os seus fluxos no processo de alimentação, gestão do armazenamento e também gestão do desempenho do sistema. De recordar que, podem ter sido estabelecidos alguns critérios em termos de desempenho que são necessários respeitar durante o dia-a-dia do *data warehouse*.

No projecto de *data warehousing* da Sonae Indústria, o planeamento do crescimento do *data warehouse* foi estimado em cada iteração para 2 anos, pretendendo-se que este armazene até 5 anos de dados históricos. A manutenção do *data warehouse* é assegurada pela equipa de desenvolvimento. De lembrar que este projecto foi desenvolvido internamente. Logo terá de ser uma equipa interna com conhecimentos de *data warehousing* a fazer a manutenção, neste caso a própria equipa de desenvolvimento.

Na Enabler, a manutenção dos projectos de *data warehousing* está dividida em duas partes. A parte referente ao processo de alimentação diário e monitorização dos seus fluxos. Esta tarefa é normalmente assegurada pelo cliente com o suporte da equipa de desenvolvimento em casos extremos. E por outro lado a gestão de infra-estruturas do *data warehouse*, nomeadamente Servidores, SGBD, Rede, etc., que também é responsabilidade do cliente, chegando ao ponto de, como se verifica no projecto de *data warehousing* da MCH, existir uma equipa especializada na arquitectura técnica. No entanto, a equipa do projecto de *data warehousing* da MCH tem já a tempo inteiro um DBA para a gestão de projectos que não incluam a equipa de gestão de infra-estruturas.

6.8 Aspectos a Salientar

No decorrer das entrevistas foram mencionados factos, tarefas e preocupações que não encaixam directamente no modelo de Kimball que foi estudado, daí que, estes sejam agora abordados e comentados.

No decorrer da entrevista com José Agostinho foi mencionada a preocupação existente no desenvolvimento de cada novo projecto ou requisito quanto ao seu impacte nos SI da Sonae Indústria. É feita uma análise extensiva a toda a área de negócio da Sonae e não apenas às necessidades do cliente local. Desta forma consegue-se manter um crescimento sustentável dos Sistemas de Informação Globais.

A existência de um *data warehouse* na Sonae Indústria tem exigido cada vez mais dos seus sistemas fonte, tendo sido necessário, inclusive, corrigir ou adicionar novos requisitos nesses sistemas, tornando-os cada vez mais exigentes a nível de introdução de dados por parte dos utilizadores.

Quer a Enabler quer a PT Inovação declaram seguir a metodologia de Kimball no desenvolvimento do *data warehouse*. Declaram também que as equipas que têm de momento estão subdimensionadas, o que tem contribuído para que, no caso da Enabler, os prazos de desenvolvimento de novos requisitos tenham sido alargados (o que por vezes não é aceitável para o cliente), e no caso da PT Inovação, tem originado a que esta tenha decidido não se candidatar a alguns concursos públicos de desenvolvimento de projectos de *data warehousing* dentro da sua especialização.

Nos projectos em que a Enabler participa, uma vez concluída a iteração em desenvolvimento, é re-analisada a ordem das iterações seguintes em conjunto com o cliente. A finalizar cada projecto é realizada uma reunião interna de fecho de projecto com o objectivo de avaliar as tarefas efectuadas no desenvolvimento do projecto, identificar tópicos de melhoria e problemas.

Dado o rápido crescimento do *data warehouse* da MCH, a Enabler viu-se obrigada a regularmente despoletar acções estruturantes, de forma a que o crescimento do *data warehouse* seja possível. Um exemplo duma dessas acções foi o desenvolvimento de um motor que controlasse as precedências de todos os processos existentes no *data warehouse*.

O responsável pela área de BI da PT Inovação, João Pias, advoga 3 máximas que um gestor de projecto deve seguir para poder ser bem sucedido:

- rodear-se de elementos competentes;
- saber delegar responsabilidades;
- responsabilizar.

Como conselho adicional, agora referente à área de *Data Warehousing*, João Pias sugeriu que periodicamente sejam efectuados controlos aos valores existentes nas fontes operacionais e aos valores existentes no *data warehouse* de forma a garantir um mínimo de consistência entre ambos.

6.9 Resumo

Analisando as entrevistas efectuadas, chega-se à conclusão que o ênfase cometido por cada uma das empresas aos projectos em que estão envolvidos é diferente. Após concluído o estudo dos projectos de *data warehousing*, definido o plano do projecto com as respectivas tarefas e identificados os pontos críticos, é possível identificar quais os aspectos de maior importância para cada uma das empresas entrevistadas.

Empresa / Projecto	Pontos Fortes
Enabler, Solutions for Retailing – Projecto de <i>Data Warehousing</i> da MCH	<ul style="list-style-type: none"> • Planeamento do Projecto • Desenvolvimento das ferramentas ETI • <i>Data warehouse</i> bastante ágil • Relacionamento com o Cliente
Enabler, Solutions for Retailing – Projecto de <i>Data Warehousing</i> da DeSpar	<ul style="list-style-type: none"> • Planeamento do Projecto • Relacionamento com o Cliente
Sonae Indústria	<ul style="list-style-type: none"> • Conhecimento do Negócio • Satisfação do Cliente
PT Inovação	<ul style="list-style-type: none"> • Gestão do Projecto • Gestão da Equipa de <i>Data Warehousing</i> • Qualidade do Projecto

Tabela 7 - Pontos Fortes dos Projectos SDW das Empresas Entrevistadas

Capítulo 7

Conclusões

7.1 Planeamento Estratégico

O desenvolvimento de um produto ou serviço requer um planeamento antecipado como forma de definir o que se quer, como se quer, qual o caminho a seguir para atingir os objectivos, e quais os recursos necessários. A este conjunto de passos denomina-se de planeamento estratégico.

Nos dias de hoje, onde a mudança e inovação são constantes, e em que o cliente é extremamente exigente e difícil de fidelizar, o planeamento estratégico do negócio de uma organização é fundamental para a sua sobrevivência e desenvolvimento.

O planeamento estratégico do negócio de uma organização visa a obtenção de vantagens sobre os seus mais directos concorrentes como forma de sobrevivência. Estas vantagens resumem-se a cinco tipos diferentes: Custo, Valor Acrescentado, Focus, Rapidez e Flexibilidade.

Exemplos de vantagens obtidas por uma organização em comparação às suas concorrentes são a comercialização de um produto ou serviço a um custo inferior, ou de qualidade superior ao dos concorrentes. Orientação da organização para a satisfação das necessidades do cliente, rapidez na execução dos pedidos do cliente e flexibilização da organização para uma melhor adaptação às mudanças ou inovações que possam acontecer no mercado.

Uma componente do planeamento estratégico do negócio é o desenvolvimento dos SI que servem de suporte à organização e que lhe permitirão atingir os objectivos definidos no plano. O planeamento associado ao desenvolvimento de SI evoluiu a par da evolução do planeamento estratégico de negócio. A evolução foi tal que o planeamento estratégico de negócio é elaborado em conjunto com o planeamento dos SI, uma vez que os SI são um recurso crítico para sucesso do planeamento estratégico.

No entanto, o planeamento e desenvolvimento de SI têm tido taxas de insucesso elevadas comprometendo assim os objectivos a atingir definidos no planeamento estratégico de negócio. Este insucesso deriva da ocorrência de uma ou mais de quatro tipos de falhas a ter em atenção no planeamento e desenvolvimento de SI, nomeadamente:

- Falha de correlação acontece quando não são atingidos os objectivos.
- Falha no processo de desenvolvimento quando este não respeita as restrições orçamentais e temporais.
- Falha de interacção entre os SI e os utilizadores.
- Falha de expectativa acontece quando as expectativas geradas pelos utilizadores não são atingidas, pondo o projecto em causa e podendo dessa forma originar uma falha de interacção.

Uma vez que o planeamento estratégico do negócio de uma organização deve estar em constante revisão com o intuito de o adaptar às constantes inovações que vão surgindo, os SI deverão também reflectir essa constante reavaliação e adaptação.

7.2 Ciclo de Vida de um Projecto de um Sistema de *Data Warehousing*

A adopção, por parte das organizações, de sistemas de *data warehousing* tem o propósito de complementar os sistemas transaccionais e conferir aos sistemas de informação da organização o potencial de análise e de armazenamento de dados históricos para os quais os sistemas transaccionais não foram desenhados.

O impacte que o desenvolvimento de um SDW tem numa organização é enorme, uma vez que vai permitir aos gestores do negócio a possibilidade de analisar os dados dos seus sistemas transaccionais numa plataforma concebida e destinada ao armazenamento de grandes volumes de informação, e que servirá de suporte à tomada de decisões.

Durante o desenvolvimento do projecto a equipa responsável deverá tomar especial atenção a factores que poderão elevar o risco do projecto. Como vimos, a correcta gestão dos factores críticos determinam o sucesso ou insucesso do projecto. Um dos principais factores críticos está relacionado com as expectativas dos utilizadores. Acima de tudo, os utilizadores deverão participar no desenvolvimento do projecto e deverão estar conscientes do objectivo do projecto, das capacidades do SDW e da sua entrada em funcionamento. A má gestão das expectativas dos futuros utilizadores do sistema poderá levar estes a alienar o sistema e condená-lo ao abandono, falhando assim um dos seus objectivos principais que é de servir de suporte à tomada de decisões.

A existência de um bom patrocinador do projecto, de preferência da área de negócio, é também um factor crítico pois permitirá garantir a rápida resolução de questões que possam perigar o projecto durante o seu desenvolvimento, sejam estas questões financeiras, temporais ou de relacionamento com departamentos da organização.

A equipa de desenvolvimento deverá ser composta por elementos com experiência e conhecimentos adequados à complexidade do projecto de forma a garantir a rápida resolução de problemas no dia-a-dia do desenvolvimento.

Existem, no entanto, outros factores de grande importância a ter em consideração durante o projecto, entre outros, por exemplo, a correcta calendarização das tarefas a realizar durante o projecto, e a elaboração de um orçamento realista que suporte o total desenvolvimento do SDW.

Tendo em conta todos estes factores, deve ser elaborado um plano do projecto a seguir durante todo o processo de criação do SDW. A tarefa de desenvolvimento monolítico de um *data warehouse* é de elevada complexidade e risco, tendo sido colocada de parte por todos aqueles que alguma vez estiveram envolvidos em projectos de *data warehousing*. Sendo assim, a construção de um SDW é um processo iterativo, uma vez que uma iteração tem por base a conclusão da iteração anterior.

Em cada iteração da construção de um SDW são identificadas 5 fases. A fase de análise que é compreendida por um conjunto de tarefas entre as quais a análise da necessidade de implementação de um SDW na organização e a elaboração do Projecto tendo como base uma análise ponderada dos requisitos orientados ao negócio. A fase planeamento, onde cabe ao Gestor do Projecto elaborar um documento com os recursos humanos, temporais e técnicos necessários à realização de cada tarefa. A fase de desenvolvimento onde são implementadas todas as tarefas que foram especificadas na fase de desenho. A fase instalação e formação onde deve ser feita a instalação de todo o *hardware* e *software* necessário à entrega do SDW, e por último, deve também ser dada formação adequada a todos os utilizadores do SDW de forma a maximizar a utilização e as potencialidades do novo sistema.

Após concluída a primeira iteração do desenvolvimento do SDW, e caso comecem a ser solicitadas novas funcionalidades ao sistema, é possível então começar uma nova iteração. De notar, que em paralelo com as próximas iterações, surgirá uma tarefa constante denominada de pós-produção que servirá de suporte aos utilizadores, e que também monitorizará o sistema e todas as tarefas ligadas ao refrescamento de dados de forma a garantir não só o nível do desempenho acordado com os utilizadores, mas também assegurar que os dados que são alimentados no sistema estejam correctos mesmo que os sistemas transaccionais sofram alterações.

Se após ter sido tomada a decisão de criação de SDW, por parte de uma organização, se tomar as precauções enunciadas neste dissertação e for feito um planeamento adequado e sistemático do projecto dividindo-o em fases e concentrando os esforços na resolução das questões de cada fase com os recursos adequados, a criação de um SDW deixará de ser um projecto de alto risco e deixará de ter as elevadas taxas de insucesso que o caracterizam.

7.3 Estudo e Análise de Diferentes Metodologias

Após a análise das diferentes perspectivas sobre o projecto, gestão e implementação de SDW de diferentes autores, chega-se à conclusão que todos consideram o processo de desenvolvimento do *data warehouse* como iterativo. Inmon defendia no início uma abordagem *Top-Down* excessivamente centrada nos dados, defendendo até o prévio desenvolvimento da base de dados do *data warehouse* para só depois abordar os respectivos temas. Tal abordagem mostrou-se incompatível com a necessidade dos dias de hoje de apresentar resultados rapidamente, daí que tenha alterado um pouco a sua visão do desenvolvimento de SDW.

É também ponto assente que diferentes tarefas podem e devem ser executadas em paralelo de forma a minimizar o tempo de desenvolvimento do projecto, e que um projecto de um SDW apenas alcançará os seus objectivos se os utilizadores a quem este se destina estiverem envolvidos no processo de planeamento e desenvolvimento. O desenvolvimento de SDW à margem dos utilizadores nunca terá a aceitação de um SDW em que a participação destes tenha sido encorajada e aproveitada.

A perspectiva de Kimball na elaboração do projecto parece ser a mais adequada, identificando quais as necessidades, tarefas, e recursos envolvidos em cada passo. No entanto, a escolha do tema a implementar na iteração em questão deveria seguir um método formal. Um exemplo de um método formal para escolha do tema, que mais impacte e melhores resultados trará à organização, é fornecido por Michael Gonzales [Gonzales '03].

Sid Adelman [Adelman et al. '00] aborda com maior profundidade a gestão do projecto de *data warehousing*, enumerando, descrevendo e propondo soluções para os diferentes tópicos nas diferentes fases que um projecto de *data warehousing* atravessa, nomeadamente a gestão de riscos, definição de objectivos, definição de critérios de avaliação e da quantificação do sucesso da implementação.

Inmon [Inmon '02] apresenta uma metodologia de desenvolvimento de sistemas de apoio à decisão mais completa, desde a metodologia usada no desenvolvimento dos sistemas operacionais, ao desenvolvimento do *data warehouse* e ferramentas que o utilizam. Em cada uma das metodologias são descritas as tarefas, a ordem em que estas são executadas, as tarefas que a precedem e sucedem assim como os resultados que são esperados e o tempo necessário à sua conclusão.

Kimball apresenta o ciclo de vida dimensional do negócio [Kimball et al. '98], identificando perfis, tarefas e responsabilidades existentes num projecto de *data warehousing*. Identifica principalmente três áreas de desenvolvimento paralelo: dados, infra-estrutura técnica e aplicações. Através do uso deste paralelismo consegue-se reduzir o tempo de entrega do projecto, mas incrementa-se os riscos e a dificuldade da tarefa de integração das três áreas no ponto de junção.

A gestão dos metadados não é abordada com a devida profundidade nas metodologias preconizadas por Kimball e Inmon. No entanto, Adelman na exposição da sua metodologia de desenvolvimento menciona a aquisição de informações referentes aos metadados para posterior uso nas ferramentas de interface com o *data warehouse*.

Desenvolvimento Iterativo do SDW.
Elaboração de um plano calendarizado com recursos técnicos e humanos necessários à execução de cada tarefa.
Método formal de selecção do próximo tema a implementar.
Desenvolvimento das fases em paralelo, sempre que possível, de forma a minimizar o tempo de conclusão.
Desenvolvimento de um plano de Comunicação.
Gestão de Metadados.
Gestão de Expectativas através do envolvimento dos utilizadores finais.
Formação dos utilizadores finais.

Tabela 8 - Tópicos Importantes na Elaboração do Projecto, na Gestão e na Implementação de SDW

7.4 Gestão do Projecto

Um planeamento cuidado e minucioso é provavelmente o factor mais importante para o desenvolvimento e conclusão com sucesso de qualquer projecto. O resultado da fase de planeamento é a elaboração de um plano calendarizado de execução do projecto subdividido em fases mais simples e de fácil gestão, atribuindo-lhes os recursos técnicos, humanos e financeiros necessários.

Cada fase é gerida através da identificação de um conjunto de processos chave que pertencem a um dos seguintes grupos de processos: grupo de processos de iniciação, planeamento, execução, controlo e encerramento. Os processos de iniciação autorizam o início da fase, os processos de planeamento definem o âmbito e tarefas a desenvolver, os processos de execução implementam o plano desenvolvido pelos processos de planeamento e os processos de controlo monitorizam a execução desse plano identificando problemas e reajustando a execução ou planeamento de forma a atingir os objectivos propostos. Os processos de encerramento dão por concluída a fase documentando os resultados obtidos.

No entanto, além dos processos estarem classificados como pertencendo a um determinado grupo de processos, estes também são relativos a uma área do conhecimento. As diferentes áreas do conhecimento abordadas na gestão de projectos reflectem as preocupações existentes durante o desenvolvimento de um projecto, são elas: Gestão da Integração, a Gestão do Âmbito, a Gestão do Tempo, a Gestão do Custo, a Gestão da Qualidade, a Gestão dos Recursos Humanos, a Gestão da Comunicação, a Gestão do Risco e a Gestão de Compras e Subcontratação.

A aplicação de todos os conhecimentos adquiridos na gestão de projectos a um projecto específico como é um projecto de SDW, é por si só uma contribuição importante para o seu sucesso, no entanto, os projectos de SDW são relativamente recentes e incorporam algumas particularidades não abordadas nos projectos comuns que elevam o grau de risco e incerteza do projecto, nomeadamente a gestão das expectativas dos *stakeholders*, a gestão dos metadados, o desenvolvimento das ferramentas de ETI, a gestão da comunicação, e gestão de recursos humanos, temporais e financeiros.

Num projecto de *data warehousing* as tarefas de planeamento e controlo devem ser mais restritivas, devendo-se prestar especial atenção às tarefas específicas do desenvolvimento de um *data warehouse*. Dado que o desenvolvimento de SDW é relativamente recente, a falta de experiência quer dos gestores de projecto quer dos elementos da equipa do projecto também contribui para o elevado o grau de risco do projecto, sendo por isso necessário um controlo mais apertado à execução de cada tarefa.

Ao contrário da maioria dos projectos, os projectos de SDW são especialmente sensíveis a um conjunto de factores que deverão ser abordados com atenção redobrada. A existência e manutenção de um patrocinador da área do negócio que viabilize o projecto, a definição e controlo do âmbito do projecto e o envolvimento dos utilizadores finais no projecto é essencial para a sua aceitação e como tal essencial ao sucesso do projecto. Em termos técnicos, factores como a qualidade dos dados dos sistemas fonte, e o desenvolvimento das ferramentas de extracção, transformação e alimentação dos dados são também importantes uma vez que o tempo que estas tarefas irão consumir é de difícil planeamento.

A finalizar, o gestor de um projecto deve estabelecer um plano de comunicação, isto é, angariação e disseminação de informação de todos e para todos os *stakeholders* do projecto, facto este ainda mais relevante nos projectos de SDW dada a sua complexidade, dimensão, custo e impacte a nível organizacional.

Foi também analisada nesta dissertação a integração da *framework* do PMBoK [PMI '00] com o ciclo de vida dimensional do negócio de Kimball [Kimball et al. '98], e conclui-se que se começa a notar uma cada vez maior preocupação com a gestão do projecto no desenvolvimento de SDW, começando mesmo a verificar-se que algumas das áreas do conhecimento e grupos de processos apresentados pela PMI aparecem contemplados na metodologia de Kimball.

Esta evolução na gestão de projecto de *data warehousing* é uma tentativa de melhorar os rácios de sucesso/insucesso no desenvolvimento de SDW e tornar o desenvolvimento de um projecto de alto risco num projecto tangível a todas as organizações que pretendam através do uso dos dados existentes nos sistemas de informação extrair conhecimento, que servirá de suporte à tomada de decisões e como tal a uma melhor adaptação à constante mudança do nosso dia-a-dia.

7.5 **Apreciação das Abordagens das Empresas**

Após entrevistadas e analisadas as metodologias empregues pelas empresas em projecto de *data warehousing*, foram esquematizadas na Figura 13 a relação existente entre o conjunto de tarefas sugeridas por Kimball em cada um das fases do projecto, representado pelo alvo, e a proximidade das tarefas executadas pelas empresas. Quanto mais perto do alvo estiver o símbolo da empresa, maior é a sua proximidade com a metodologia de Kimball para essa fase.

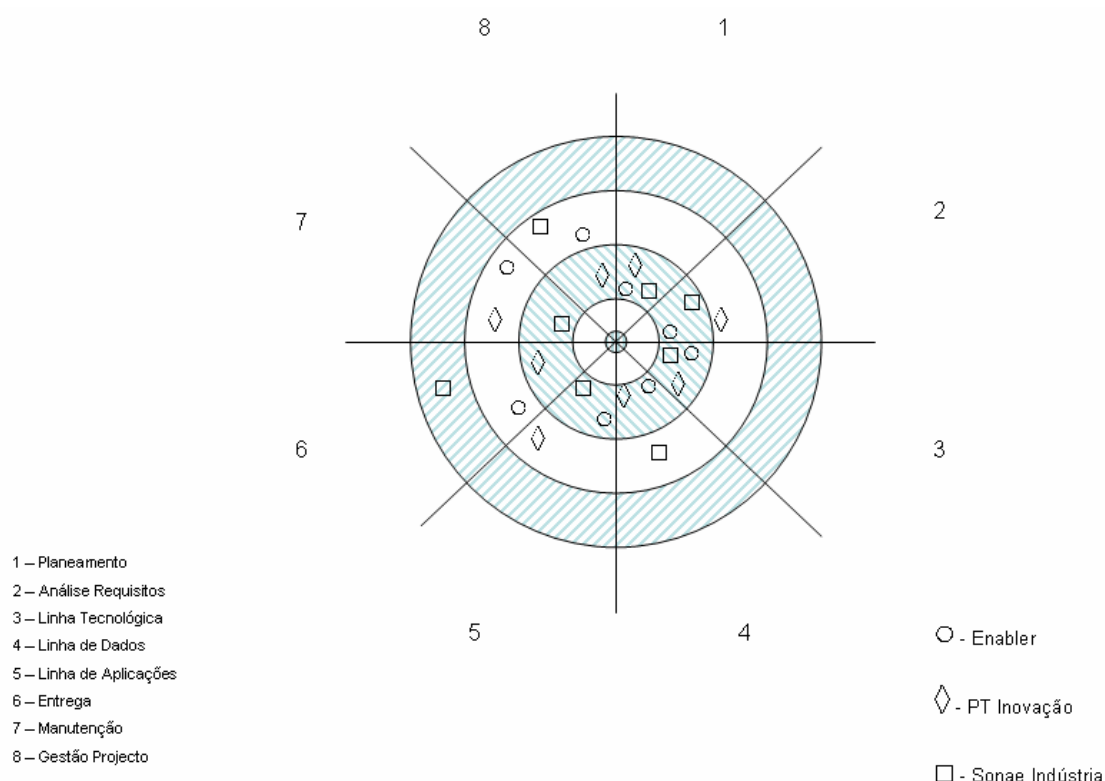


Figura 11 - Áreas Alvo

A Figura permite-nos assim concluir que as tarefas associadas às fases de planeamento e análise requisitos são executadas pelas três empresas de forma semelhante e em consonância com a metodologia de Kimball. Tal, deve-se ao facto de serem fases comuns ao desenvolvimento de um qualquer projecto de *software* e não apenas ao desenvolvimento de Sistemas de *Data Warehousing*. As tarefas inerentes à fase da Linha de Acção de Dados são executadas pela Enabler e pela PT Inovação de acordo com a metodologia de Kimball. De salientar que a Sonae Indústria não segue o modelo dimensional sugerido por Kimball para o desenho do seu *data warehouse*. Na fase de Entrega, nota-se alguma dificuldade no cumprimento dos prazos por parte da Sonae Industria. Tal deve-se à dificuldade da gestão do âmbito do projecto uma vez que o SDW foi desenvolvido internamente. Por outro lado, este factor torna-se uma vantagem na análise das tarefas de manutenção do *data warehouse*. A nível de gestão de projecto, é de salientar a estrutura adoptada pela PT Inovação para a gestão do projecto e dos elementos que o constituem.

7.6 Apontamento Final

Esta dissertação teve como principal objectivo o estudo do desenvolvimento de projectos de *data warehousing*, analisando a envolvente histórica que levou ao seu aparecimento como um componente de um SI de uma organização. Foi também analisado o ciclo de vida de um projecto de *data warehousing*, o impacte da sua implementação na organização, quais as tarefas existentes no projecto e os perfis adequados para as executar, assim como, foram identificados alguns factores críticos (gestão da expectativas, o envolvimento dos utilizadores, bom patrocinador, equipa técnica competente, calendarização, orçamentação) que determinam o sucesso ou insucesso do projecto. Em seguida, foram analisadas e comparadas duas metodologias de autores reputados na área de *data warehousing*, Ralph Kimball e William H. Inmon, e complementadas com opiniões de outros autores em aspectos menos claros. A finalizar foram efectuadas algumas entrevistas a empresas que estão envolvidas em projectos de *data warehousing*, nas quais o objectivo fundamental era conhecer a sua metodologia de gestão e desenvolvimento dos projectos, comparando-a sempre que possível às metodologias estudadas previamente.

Como última nota, deve-se referir que os projectos de desenvolvimento de Sistemas de *Data Warehousing* são parecidos, em forma, com qualquer outro projecto de desenvolvimento de outro tipo de sistema. No entanto, os SDW têm particularidades técnicas (p.ex.: a extracção, a transformação e a integração de dados) que exigem um planeamento e desenvolvimento bastante cuidado. Este facto não é, por si só, responsável pela grande percentagem de insucesso no desenvolvimento de SDW, uma vez que, outro tipo de projectos têm também as suas especificidades técnicas. O cerne da questão é, e continuará a ser, a Gestão do Projecto. O desenvolvimento ou adopção de uma metodologia aplicada à Gestão de Projectos contribuirá de forma significativa para a redução do insucesso dos projectos de *data warehousing*, o que resultará, naturalmente num aumento das intenções de desenvolvimento de SDW a nível nacional, impulsionando consequentemente o desenvolvimento da área.

Bibliografia

[Adelakun '96] O. Adelakun: "Quality Planning: A Conceptual Framework for Improving Information Systems Success", Exploring the Limits of Support Systems, 1-16, December 1996, 1996

[Adelman '01] S. Adelman: "Ten Worst Practices of the Unsucessful Data Warehouse Project Manager", http://doe.k12.hi.us/technology/busintel/sa_dw10worstpractices.htm, Acedido em 15-01-2004

[Adelman et al. '00] S. Adelman e L. T. Moss: "Data Warehouse Project Management", 1st Ed., I. Technology, Addison-Wesley, 2000

[Adelman et al. '98] S. Adelman e J. Oates: "Data Warehouse Project Management", http://www.dmreview.com/editorial/dmreview/print_action.cfm?EdID=604, Acedido em 15-01-2004

[ADS '95] ADS: "data warehousing survey", The Meta Group Inc.

[Agosta '00] L. Agosta: "The Essential Guide to Data Warehousing", 1st Ed., Prentice Hall, 2000

[Allen '95] W. E. Allen: "Establishing some basic project-management body-of-knowledge concepts", *International Journal of Project Management* 13 (2), 1995, 77-82

[APM '00a] APM: "Project Management : Body of Knowledge", Fourth Ed., Association for Project Management, 2000a

[APM '00b] APM: "Syllabus for the APMP Examination", Second Ed., Association for Project Management, 2000b

[Armstrong '01] R. Armstrong: "Seven Steps to Optimizing Data Warehouse Performance", *Computer IEEE* 2001, 76-79

[Ayas '95] K. Ayas: "Professional project management: a shift towards learning and a knowledge creating structure", *International Journal of Project Management* 14 (3), 1995, 131-136

[Blackwood '00] P. Blackwood: "11 Steps To Successful Data Warehousing", <http://www.advisor.com/Articles.nsf/dp/EBACCDC99D0DC471852568D2004B968A>, Acedido em 25-06-2003

[Boar '97] B. H. Boar: "Building, Using, and Managing the Data Warehouse-Understanding Data Warehousing Strategically", 1st Ed., Prentice Hall PTR, 1997

[Boehm '88] B. W. Boehm: "A Spiral Model of Software Development and Enhancement", *Computer IEEE* 1988, 61-72

[Bonifati et al. '01] A. Bonifati, F. Cattaneo, S. Ceri, A. Fuggeta e S. Paraboschi: "Designing Data Marts for Data Warehouses", *ACM Transactions on Software Engineering and Methodology* 10 (4), 2001, 452-483

[Bontempo et al. '98] C. Bontempo e G. Zagelow: "The IBM Data Warehouse Architecture", Communications of the ACM 41 (9), 1998, 38-48

[Cerpa et al. '98] N. Cerpa e J. M. Verner: "Case study: The effect of IS maturity on information systems strategic planning", Information & Management 34 1998, 199-208

[Chan et al. '98] Y. E. Chan, S. L. Huff e D. G. Copeland: "Assessing realized information systems strategy", Journal of Strategic Information Systems 6 1998, 273-298

[Chaudhuri et al. '96] S. Chaudhuri e U. Dayal: "An Overview of Data Warehousing and OLAP Technology", VLDB Conference, 1996

[Clarke et al. '97] S. Clarke e B. Lehaney: "Information Systems Strategic Planning: A Model for Implementation in Changing Organizations", Syst. Res. Behav. Sci. 14 (2), 1997, 129-136

[Cleland '95] D. I. Cleland: "Leadership and the project-management body-of-knowledge", International Journal of Project Management 13 (2), 1995, 83-88

[Connolly et al. '02] T. Connolly e C. Begg: "Database Systems - A Practical Approach to Design, Implementation, and Management", Third Ed., 2002

[Cooper et al. '00] B. L. Cooper, H. J. Watson, B. H. Wixom e D. L. Goodhue: "Data Warehousing Supports Corporate Strategy at First American Corporation", MIS Quarterly 24 (4), 2000, 547-567

[Craig et al. '99] R. S. Craig e J. A. Vivona: "Microsoft Data Warehousing: Building Distributed Decision Support Systems", Ed., John Wiley & Sons, 1999

[Duncan '95] W. R. Duncan: "Developing a project-management body-of-knowledge document: the US Project Management Institute's approach, 1983-1994", International Journal of Project Management 13 (2), 1995, 89-94

[Dutta et al. '95] S. Dutta e Y. Doz: "Linking information technology to business strategy at Banco Comercial Português", *Journal of Strategic Information Systems* 4 (1), 1995, 89-110

[Faber et al. '02] N. Faber, R. B. M. Koster e S. L. van de Velde: "Linking warehouse complexity to warehouse planning and control structure", *International Journal of Physical Distribution & Logistics Management* 32 (5), 2002, 381-395

[Fleming et al. '98] Q. W. Fleming e J. M. Koppelman: "Earned Value Project Management : A Powerful Tool for Software Projects", *Crosstalk : The Journal of Defense Software Engineering* 1998, 19-23

[Flowers '96] S. Flowers: "Software failure: management failure", Ed., John Wiley, 1996

[Foster '97] K. Foster: "Information Delivery : A Project Management Approach", 22nd Annual SAS Users Group International Conference, San Diego, California, 635-640, 1997

[Franconi et al. '03] E. Franconi e A. Kamble: "The GMD Data Model for Multidimensional Information: a brief introduction", 5th International Conference on Data Warehousing and Knowledge Discovery, Prague, 2003

[Franconi et al. '99] E. Franconi e U. Sattler: "A Data Warehouse Conceptual Model for Multidimensional Aggregation", International Workshop on Design and Management of Data Warehouses, Heidelberg, Germany, 1999

[Furrow '01] G. Furrow: "The Case for Building a Data Warehouse", *IT Pro IEEE* 2001, 31-33

[Gallas '99] S. Gallas: "Kimball Vs. Inmon", *DM Direct*, September 1999

[Gardner '98] S. R. Gardner: "Building the Data Warehouse", *Communications of the ACM* 41 (9), 1998, 52-60

[Gatzui et al. '99] S. Gatzui, M. A. Jeusfeld, M. Staudt e Y. Vassiliou: "Design and Management of Data Warehouses Report on the DMDW'99 Workshop", 1999

[Golfarelli et al. '98] M. Golfarelli e S. Rizzi: "A Methodological Framework for Data Warehouse Design", ACM First International Workshop on Data Warehousing and OLAP, Washington D.C., USA, 1998

[Golfarelli et al. '99] M. Golfarelli e S. Rizzi: "Designing the Data Warehouse: Key Steps and Crucial Issues", Computer Science and Information Management 2 (3), 1999, 1-14

[Gonzales '03] M. L. Gonzales: "IBM Data Warehousing", Ed., John Wiley & Sons, 2003

[Gottschalk et al. '01] P. Gottschalk e H. Solli-Saether: "Integration between Business Planning and Information Systems Planning: An Analysis of Technology Exploration and Exploitation in Different Value Configurations", 34th Hawaii International Conference on Systems Science, 9, 2001

[Gray et al. '98] P. Gray e H. J. Watson: "Present and Future Directions in Data Warehousing", The DATABASE for Advances in Information Systems 29 (3), 1998, 83-90

[Greenfield '02a] L. Greenfield: "The Case Against Data Warehousing", <http://www.dwinfocenter.org/against.html>, Acedido em 14-06-2003

[Greenfield '02b] L. Greenfield: "The Case for Data Warehousing", <http://www.dwinfocenter.org/casefor.html>, Acedido em 13-06-2003

[Greenfield '02c] L. Greenfield: "An (Informal) Taxonomy of Data Warehouse Data Errors", <http://www.dwinfocenter.org/errors.html>, Acedido em 13-06-2003

[Grundy '98] T. Grundy: "Strategy implementation and project management", International Journal of Project Management 16 (1), 1998, 43-50

[Gupta '97] V. R. Gupta: An Introduction to Data Warehousing", Ed., 1997

[Gyssens et al. '97] M. Gyssens e L. V. S. Lakshmanan: "A Foundation for Multi-Dimensional Databases", 23rd VLDB Conference, Athens, Greece, 106-115, 1997

[Hammer et al. '95] J. Hammer, H. Garcia-Molina, W. Jennifer, W. Labio e Y. Zhuge: "The Stanford Data Warehousing Project", IEEE Data Engineering Bulletin 18 (2), 1995, 41-48

[Herdlein '96] S. Herdlein: "Data Warehouse Project Management: How to Build a Solid Foundation For Your Data Warehouse", <http://www.tdan.com/i002fe08.htm>, Acedido em 22-11-2003

[Hernández et al. '95] M. A. Hernández e S. J. Stolfo: "The Merge/Purge Problem for Large Data Databases", 127-138, 1995

[Hernández et al. '98] M. A. Hernández e S. J. Stolfo: "Real-world Data is Dirty: Data Cleansing and The Merge/Purge Problem", 1998,

[Hevner et al. '00] A. R. Hevner, D. J. Berndt e J. Studnicki: "Strategic Information Systems Planning with Box Structures", 33rd Hawaii International Conference on System Sciences, Hawaii, 2000

[Hidding '01] G. J. Hidding: "Sustaining strategic IT advantage in the information age: how strategy paradigms differ by speed", Journal of Strategic Information Systems 10 2001, 201-222

[Huseman et al. '00] B. Huseman, J. Lechtenborger e G. Vossen: "Conceptual Data Warehouse Design", International Workshop on Design and Management of Data Warehouses, Stockholm, Sweden, 2000

[Huynh et al. '00] T. N. Huynh, O. Mangisengi e A. M. Tjoa: "Metadata for Object-Relational Data Warehouse", International Workshop on Design and Management of Data Warehouses, Stockholm, Sweden, June 5-6, 2000

[Imhoff '00] C. Imhoff: "The Corporate Information Factory", <http://www.datawarehouse.com/iknowledge/articles/print.cfm?ContentID=145>, Acedido em 22-11-2003

[Inmon '94] W. H. Inmon: "Using the Data Warehouse", 1st Ed., John Wiley & Sons, 1994

[Inmon '97] W. H. Inmon: "Managing the Data Warehouse", 1st Ed., John Wiley & Sons, 1997

[Inmon '00a] W. H. Inmon: "Data Mart Does Not Equal Data Warehouse", <http://www.datawarehouse.com/iknowledge/articles/print.cfm?ContentID=155>, Acedido em 14-07-2003

[Inmon '00b] W. H. Inmon: "A Data Warehouse Development Methodology", <http://www.billinmon.com/library/methods/dwmeth.asp>, Acedido em 22-11-2003

[Inmon '00c] W. H. Inmon: "The Future of Data Warehousing: Alternative Storage", <http://www.filetek.com/papers/Inmon/inmon.htm>, Acedido em 26-06-2003

[Inmon '02] W. H. Inmon: "Building the Data Warehouse", 3rd Ed., W. C. Publishing, John Wiley & Sons, 2002

[Inmon et al. '01] W. H. Inmon e D. P. Meers: "Accelerating Enterprise Data Warehouse Profitability: A New Approach to Design and Deployment", 2001,

[Inmon et al.] W. H. Inmon e R. Sousa: "The Metadata Component",

[Jarke et al. '98] M. Jarke, M. Lenzerini, Y. Vassiliou e P. Vassiliadis: "Fundamentals of Data Warehouses", Ed., Springer, 1998

[Jeusfeld et al. '98] M. A. Jeusfeld, C. Quix e M. Jarke: "Design and Analysis of Quality Information for Data Warehouses", 17th International Conference on Conceptual Modeling / the Entity Relationship Approach, Singapore, 349-362, November 1998, 1998

[Jones '98] K. Jones: "An Introduction to Data Warehousing: What Are the Implications for the Network", International Journal of Network Management 8 1998, 42-56

[Jordan et al. '95] E. Jordan e B. Tricker: "Information Strategy: alignment with organization structure", Journal of Strategic Information Systems 4 (4), 1995, 357-382

[Kearns et al. '00] G. S. Kearns e A. L. Lederer: "The effect of strategic alignment on the use of IS-based resources for competitive advantage", Journal of Strategic Information Systems (9), 2000, 265-293

[Kimball et al. '98] R. Kimball, L. Reeves, M. Ross e W. Thornthwaite: "The Data Warehouse Lifecycle Toolkit - Expert Methods for Designing, Developing, and Deploying Data Warehouses", Ed., John Wiley & Sons, Inc., 1998

[Kimball et al. '02] R. Kimball e M. Ross: "The Data Warehouse Toolkit - The Complete Guide to Dimensional Modeling", Second Ed., W. C. Publishing, John Wiley & Sons, 2002

[Lawyer et al. '04] J. Lawyer e S. Chowdhury: "Best Practices in Data Warehousing to Support Business Initiatives and Needs", 37th Hawaii International Conference on System Sciences, Hawaii, 9, 2004

[Lederer et al. '96] A. L. Lederer e H. Salmela: "Toward a theory of strategic information systems planning", Journal of Strategic Information Systems 5 1996, 237-253

[Lederer et al. '88] A. L. Lederer e V. Sethi: "The implementation of strategic information systems planning methodologies", MIS Quarterly 12 (3), 1988, 444-461

[Lee et al. '03] G.-G. Lee e R.-J. Bai: "Organizational mechanisms for sucessful IS/IT strategic planning in the digital era", Management Decision 41 (1), 2003, 32-42

[Lee et al. '99] M. L. Lee, H. Lu, T. W. Ling e Y. T. Ko: "Cleansing Data for Mining and Warehousing", 10th International Conference on Database and Expert Systems Applications, Florence, 751-760, August 1999, 1999

[Levy et al. '00] M. Levy e P. Powell: "Information systems strategy for small and medium sized enterprises : an organisational perspective", Journal of Strategic Information Systems 9 2000, 63-84

[Little Jr. et al. '99] R. G. Little Jr. e M. L. Gibson: "Identification of Factors Affecting the Implementation of Data Warehousing", 32nd Hawaii International Conference on System Sciences, Hawaii, 1-10, 1999

[Lyytinen et al. '87] K. Lyytinen e R. Hirschheim: "Information failures - a survey and classification of the empirical literature." Oxford Surveys in Information Technology 4 1987, 257-309

[Maletic et al. '00] J. I. Maletic e A. Marcus: "Data Cleansing: Beyond Integrity Analysis", Conference on Information Quality, Boston, 200-209, October 2000, 2000

[Marco '01] D. Marco: "Top Mistakes To Avoid When Building a Data Warehouse", <http://www.tdan.com/i016fe04.htm>, Acedido em 22-11-2003

[Mentzas '97] G. Mentzas: "Implementing an IS strategy - a team approach", Long Range Planning 10 (1), 1997, 84-95

[Min et al. '99] S. K. Min, E. H. Suh e S. Y. Kim: "An integrated approach toward strategic information systems planning", *Journal of Strategic Information Systems* (8), 1999, 373-394

[Moss et al. '99a] L. T. Moss e S. Adelman: "Data Warehouse Goals and Objectives, Part 1", http://www.dmreview.com/editorial/dmreview/print_action.cfm?EdID=1365, Acedido em 26-06-2003

[Moss et al. '99b] L. T. Moss e S. Adelman: "Data Warehouse Goals and Objectives: Part 2: Short-Term Objectives", http://www.dmreview.com/editorial/dmreview/print_action.cfm?EdID=1456, Acedido em 26-06-2003

[Moss et al. '99c] L. T. Moss e S. Adelman: "Data Warehouse Goals and Objectives: Part 3: Long-Term Objectives", http://www.dmreview.com/editorial/dmreview/print_action.cfm?EdID=1564, Acedido em 26-06-2003

[Munns et al. '96] A. K. Munns e B. F. Bjeirmi: "The role of project management in achieving project success", *International Journal of Project Management* 14 (2), 1996, 81-87

[Newkirk et al. '03] H. E. Newkirk, A. L. Lederer e C. Srinivasan: "Strategic information systems planning: too little or too much?" *Journal of Strategic Information Systems* 12 2003, 201-228

[Orr '00] K. Orr: "Data Warehousing Technology", <http://www.kenorrinst.com/dwpaper.html>, Acedido em 22-11-2003

[Peppard et al. '04] J. Peppard e J. Ward: "Beyond strategic information systems : towards an IS capability", *Journal of Strategic Information Systems* 2004,

[Peters et al. '02] S. C. A. Peters, M. S. H. Heng e R. Vet: "Formation of the information systems strategy in a global financial services company", *Information and Organization* (12), 2002, 19-38

[PMI '00] PMI: "A Guide to the Project Management Body of Knowledge", Ed., Project Management Institute, 2000

[Poe et al. '98] V. Poe, P. Klauer e S. Brobst: "Building a Data Warehouse for Decision Support", 2nd Ed., Prentice Hall, 1998

[Porter et al. '85] M. E. Porter e V. E. Millar: "How Information Gives You Competitive Advantage", *Harvard Business Review*, 1985

[Ravat et al. '99] F. Ravat, O. Teste e G. Zurfluh: "Towards Data Warehouse Design", *Conference on Information and Knowledge Management*, Kansas City, 359-366, 1999

[Rundensteiner et al. '00] E. A. Rundensteiner, A. Koeller e X. Zhang: "Maintaining Data Warehouses Over Changing Information Sources", *Communications of the ACM* 43 (6), 2000, 57-62

[Ruohonen '96] M. Ruohonen: "Information Technology Mediated Activities in Organizational Contexts - A Case of Strategic Information Systems Planning", 1996

[Sakaguchi et al. '96] T. Sakaguchi e M. N. Frolick: "A Review of the Data Warehouse Literature", 1996

[Salmela et al. '02] H. Salmela e T. A. N. Spil: "Dynamic and emergent information systems strategy formulation and implementation", *International Journal of Information Management* 22 2002, 441-460

[Sammon et al. '00] D. Sammon e P. Finnegan: "The Ten Commandments of Data Warehousing", *The DATABASE for Advances in Information Systems* 31 (4), 2000, 82-91

[Sen et al. '98] A. Sen e V. S. Jacob: "Industrial Strength Data Warehousing", *Communications of the ACM*, September 1998

[Shanks '97] G. Shanks: "The challenges of strategic data planning in practice: an interpretive case study", *Journal of Strategic Information Systems* 6 1997, 69-90

[Sharma '01] R. Sharma: "Strategy for Rearchitecting an Existing Data Warehouse", <http://www.datawarehouse.com/iknowledge/articles/print.cfm?ContentID=4403>, Acedido em 17-01-2004

[Shenhar et al. '96] A. J. Shenhar e D. Dvir: "Toward a typological theory of project management", *Research Policy* 25 1996, 607-632

[Shim et al. '02] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda e C. Carlsson: "Past, present, and future of decision support technology", *Decision Support Systems* 33 2002, 111-126

[Shin '02] B. Shin: "A case of data warehousing project management", *Information & Management* 39 (39), 2002, 581-592

[Soares '02] J. Soares: "Soluções de Data Warehousing - Fundamentos Teóricos, Metodologias e Práticas de Implementação", *Gestão de Sistemas de Informação*, Instituto de Ciências do Trabalho e da Empresa, Lisboa 2002

[Soderlund '04] J. Soderlund: "Building theories of project management: past research, questions for the future", *International Journal of Project Management* 22 2004, 183-191

[Srivastava et al. '99] J. Srivastava e P.-Y. Chen: "Warehouse Creation - A Potential Roadblock to Data Warehousing", IEEE Transactions on Knowledge and Data Engineering 11 (1), 1999, 118-126

[Subramanian et al. '96] A. Subramanian, L. D. Smith e A. C. Nelson: "Strategic Planning for Data Warehousing in the Public Sector", 29th Annual Hawaii International Conference on System Sciences, Hawaii, 54-60, 1996

[Subramanian et al. '97] A. Subramanian, L. D. Smith, A. C. Nelson, J. F. Campbell e D. A. Bird: "Strategic planning for data warehousing", Information & Management 33 (33), 1997, 99-113

[Sutter '98] J. R. Sutter: "Project-Based Warehouses", Communications of the ACM, September 1998

[Theodoratos et al. '00] D. Theodoratos e M. Bouzeghoub: "A General Framework for the View Selection Problem for Data Warehouse Design and Evolution", 2000

[Wateridge '97] J. Wateridge: "Training for IS/IT project managers: a way forward", International Journal of Project Management 15 (5), 1997, 283-288

[Watson et al. '01] H. J. Watson, D. A. Annino, B. H. Wixom, K. L. Avery e M. Rutherford: "Current Practices in Data Warehousing", Data Warehousing Today 2001, 47-55

[Watson et al. '98] H. J. Watson e B. J. Haley: "Managerial Considerations", Communications of the ACM, September 1998

[Westerman '01] P. Westerman: "Data Warehousing - Using the Wal-Mart Model", Ed., Morgan Kaufmann Publishers, 2001

[Wideman '95] R. M. Wideman: "Criteria for a project-management body of knowledge", International Journal of Project Management 13 (2), 1995, 71-75

[Widow '95] J. Widow: "Research Problems in Data Warehousing", 4th Internacional Conference on Information and Knowledge Management, November, 1995

[Willis '95] B. E. Willis: "APM project-management body of knowledge: the European view", International Journal of Project Management 13 (2), 1995, 95-98

[Winter '01] R. Winter: "The Current and Future Role of Data Warehousing in Corporate Application Architecture", 34th Hawaii International Conference on System Sciences, Hawaii, 7, 2001

[Winter et al. '03] R. Winter e B. Strauch: "A Method for Demand-driven Information Requirements Analysis in Data Warehousing Projects", 36th Hawaii International Conference on System Sciences, Hawaii, 2003

[Wixom et al. '00] B. H. Wixom e H. J. Watson: "An Empirical Investigation of the Factors Affecting Data Warehousing Success", MISQuarterly, March 2001 2000

[Yeo '02] K. T. Yeo: "Critical failure factors in information system projects", International Journal of Project Management 20 (20), 2002, 241-246

Referências WWW

- [1] <http://www.dwinfocenter.org/>

Site desenvolvido por Larry Greenfield, consultor da LGI Systems, com o objectivo de prestar alguns esclarecimentos sobre área de *Data Warehousing*.

- [2] <http://www-db.stanford.edu/warehousing/warehouse.html>

O projecto *WareHouse Information Prototype at Stanford* (WHIPS) teve como objectivo primordial analisar a criação e a manutenção de SDW e desenvolver algoritmos e ferramentas que assegurem estas actividades. Aqui encontra-se toda a documentação disponível sobre este projecto.

- [3] <http://www.kimballuniversity.com>

Ralph Kimball fundou a Kimball University com o intuito de fornecer cursos de modelação dimensional aplicada aos sistemas de *data warehousing*, baseando-se nos livros que tem editado nos últimos anos.

- [4] <http://datawarehouse.ittoolbox.com>

Portal dedicado ao mundo de *Data Warehousing*. Contém Blogs, FAQs, White Papers, anúncios de novidades comerciais, anúncios de empregos, etc

- [5] <http://www.inmoncif.com>

Site da Inmon Associates, Inc onde é apresentada a Corporate Information Factory (CIF) juntamente com alguns artigos de opinião e white papers de consultores especializados em *Data Warehousing*.

- [6] <http://www.dw-institute.com>

Data Warehouse Institute (TDWI) é dos principais sites de pesquisa e educação sobre *Business Intelligence* e *Data Warehousing*. Promove e patrocina conferências, seminários e cursos sobre os temas da especialidade. Tem disponíveis no site *White Papers*, e publica trimestralmente o *Business Intelligence Journal*.

- [7] <http://www.dm-review.com>

Publicação online da DM Review, uma das mais conceituadas revistas de *Business Intelligence* e *Data Warehousing*. Além dos artigos publicados na revista possui *White Papers*, artigos de opinião de autores conceituados e portais sobre áreas específicas de *Data Warehousing*.

- [8] <http://www.datawarehouse.com>

Site desenvolvido pela DM Review com informações referentes a conferências e exposições sobre *data warehousing*. Contém também fóruns de variados assuntos e acesso a artigos, relatórios técnicos, eventos e seminários na área de *Data Warehousing*.

- [9] <http://www.tdan.com>

A *Data Administration Newsletter* é, como o nome indica, uma publicação trimestral da área de administração de dados que recebe e publica artigos relacionados com o seu domínio.

- [10] <http://www.advisor.com>

A Advisor é uma organização dedicada à publicação de revistas, à organização de eventos, conferências e seminários sobre os mais variados temas de Tecnologias da Informação.

- [11] <http://www.ondelette.com/OLAP/dwbib.html>

Página internet com referências bibliográficas sobre *Data Warehousing* e OLAP.